Ab initio prediction of crystalline phases and electronic properties of alloys and other compounds

Rafael Sarmiento Perez

N° d'ordre 155−2015 Année 2015

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE DE PHYSIQUE ET ASTROPHYSIQUE

DIPLOME DE DOCTORAT
(arrêté du 7 aout 2006)

soutenue publiquement le 24 septembre 2015

par

Rafael A. Sarmiento Pérez

AB INITIO PREDICTION OF CRYSTALLINE PHASES AND ELECTRONIC PROPERTIES OF ALLOYS AND OTHER COMPOUNDS

Directeur de thèse: Dr Patrice Melinon
Co-directeur de thèse: Dr Silvana Botti

JURY

M. Alfonso San Miguel	Président du Jury
M. Razvan Caracas	Rapporteur
M. Xavier Blase	Rapporteur
M. Matteo Calandra	Examinateur
M. Tristan Albaret	Examinateur
Mme. Silvana Botti	Co-directeur de thèse
M. Patrice Melinon	Directeur de thèse

Abstract

In this work we present an *ab initio* materials design study of several systems covering intermetallic and semiconducting alloys, transparent conductive oxides and molecular solids. We performed Minima Hopping calculations combined with Density Functional Theory that made possible to unveil several stable compounds in the phase diagrams of lithium-aluminium and sodium-gold binary alloys, as well as low-symmetry geometries of $CuBO_2$, significantly lower in energy than the controversial delafossite structure reported as its ground state. We also found that the H_3 molecule can be stabilized inside Cl cages at pressures of around 100 GPa. Additionally, we combined high-throughput techniques and global structure prediction methods to find nitride perovskites structures. In a different line, we studied the change in the absorption properties of the $Cu(In,Ga)S_2$ chalcopyrite alloys as it was unexpectedly observed in experiment that with the change of the In/Ga ratio, the S K-absorption edge shifts, while the absorption edges of the other species is largely independent of the composition. In a more fundamental chapter, we propose a semiempirical exchange-correlation functional optimized to yield accurate energies of formation of solids. The manuscript is organized as follows.

Chapter 1 consists of an overview of concepts that are useful for the reading of the thesis. First, we give a short introduction to the Kohn-Sham Density Functional Theory and exchange-correlation functionals. In the following sections we describe different approaches to materials design with a focus on the Minima Hopping method, which was used for the structure prediction calculations performed in this work. Finally, the construction of the convex hull of phase stability, in particular the bidimensional case, is explained in detail as it is often used on the following chapters.

In Chapter 2 we present a methodology for the computation of binary phase diagrams from first-principles, based on the Minima Hopping method and its application for intermetallic binary alloys sodium-gold and lithium-aluminium, and for crystals of H-Cl at moderate pressures. We first consider the sodium-gold system. We obtain that the most stable composition is $NaAu_2$, in agreement with available experimental data. We also confirm the crystal structures of $NaAu_2$ and Na_2Au, that were fully characterized in experiments, and identify a candidate ground-state structure for the experimental stoichiometry NaAu. Moreover, we obtain three other stoichiometries, namely Na_3Au_2, Na_3Au, and Na_5Au, that turn out to be thermodynamically stable in

our simulations. Finally, we perform phonon calculations to check the stability of all reported phases and we simulate X-ray diffraction spectra for comparison with future experimental data.

In the section that follows, we apply the same methodology to the intermetallic Li-Al alloys. These compounds are on one hand key materials for light-weight engineering, and on the other, they have been proposed for high-capacity electrodes for Li batteries. We determine from first-principles the phase diagram of Li-Al binary crystals using the Minima Hopping method. Beside reproducing the experimentally reported phases (LiAl, Li_3Al_2, Li_9Al_4, $LiAl_3$, and Li_2Al), a structural variety larger than expected is unveiled by discovering six non-reported binary phases, likely to be thermodynamically stable. We then compare the results using different exchange-correlation functionals and we discuss the behavior of the elastic constants of all found Li-Al stable binaries as a function of their stoichiometry.

Finally, we compute the phase diagram of H-Cl binary crystals. In this case we focus on the formation of triatomic molecules of H in such crystals. We show that H_3 molecules, which are expected to become stable in pure hydrogen above 2.1 TPa, can be stabilized in H-Cl crystals at pressures about 100 GPa, perfectly achievable nowadays in any laboratory.

In Chapter 3 we present a study of materials of interest for applications in photovoltaics: the transparent conductive oxide $CuBO_2$, novel nitride perovskites with composition ABN_3 and $Cu(In,Ga)S_2$ chalcopyrite alloys. First, with our Minima Hopping structure prediction calculations, we prove that the delafossite structure of $CuBO_2$ reported experimentally, is very unlikely to be the ground state of this material as we found several other crystal structures lower in energy up to 600 meV.

In the following section we present our results on the existence of perovskites with composition ABN_3. Our approach is based on a combination of high-throughput techniques and global structure prediction methods. We find 21 new compositions of the form ABN_3 that are thermodynamically stable and that have therefore excellent chances of being experimentally accessible. Most of these materials crystallize in monoclinic phases, but three compounds, namely $LaReN_3$, $LaWN_3$ and $YReN_3$ are predicted to have distorted perovskite structures in their ground state.

The last section of the chapter presents the Density Functional Theory study of the element-specific unoccupied electronic states of $Cu(In,Ga)S_2$ as a function of the In/Ga ratio. Near the Edge X-Ray Absorption Fine Structure measurements performed on $CuIn_xGa_{1-x}S_2$ thin films for x=0, 0.67 and 1 by our experimental collaborators, show that the S absorption edge shifts with changing In/Ga ratio as expected from

the variation of the band gap. On the other hand, the Ga, In, and Cu absorption edge positions remain nearly unaltered with alloy composition despite a significant change of the band gap. Our calculations indicate that this behavior originates from the dependence of the electronic states on the local atomic environment, while the change in band gap arises from the spatial average of these localized states.

In Chapter 4 we develop a semi-empirical exchange-correlation functional for Density Functional Theory tailored to calculate energies of formation of solids. This functional has the form of the Perdew-Burke-Ernzerhof functional, but with three parameters, covering the exchange and correlation parts, fitted to reproduce experimental energies of formation for a representative set of binary compounds. The quality of the obtained functional was then assessed for a control set. Our functional manages to reduce the error of the Perdew-Burke-Ernzerhof generalized gradient approximation by roughly a factor of two. Furthermore, this is achieved without compromising the quality of the geometry.

The results collected in this manuscript represent significant examples of the level of maturity achieved at present by the domain of materials design. The variety of our contributions shows that our powerful approach to materials design is very general, and the developments brought by this thesis are now available to the whole community for many unexplored field of applications.

Résumé

Cette thèse présente une étude dans le cadre de la conception *ab initio* de nouveaux matériaux, avec des applications aux alliages intermétalliques et semi-conducteurs, aux oxydes transparents conducteurs et aux solides moléculaires. Des simulations avec la méthode Minima Hopping combinée avec la théorie de la fonctionnelle de la densité ont été utilisées pour trouver des nouveaux composés dans les diagrammes de phase des composés binaires de Lithium-Aluminium et Sodium-Or, aussi bien que des géométries de faible symétrie de $CuBO_2$ à plus basses énergies que la structure delafossite qui etait considérée comme son état fondamental. Nous avons aussi couplé la méthode Minima Hopping et la recherche de structures avec prototypes pour trouver des nouvelles perovskites de nitrure. Également, nous avons trouvé que la molécule H_3 peut être stabilisée dans des structures à cages de Cl à pressions d'environ 100 GPa. Nous avons aussi étudié les propriétés électroniques des alliages de chalcopyrite $Cu(In,Ga)S_2$. Dans un sujet plus fondamental, nous proposons une fonctionnelle d'échange-corrélation semi-empirique optimisée pour obtenir des énergies de formation plus précises pour les solides. Dans la suite nous résumons le contenu par chapitre de ce manuscrit.

Le premier chapitre donne un aperçu des concepts théoriques utiles pour la lecture de la thèse. D'abord, nous présentons une brève introduction à la théorie de la fonctionnelle de la densité et aux fonctionnelles d'échange-corrélation. Nous décrivons ensuite différentes approches de la conception *ab initio* de nouveaux matériaux et notamment la méthode Minima Hopping. Pour conclure, nous discutons le concept d'enveloppe convexe pour décrire la stabilité thermodynamique, en particulier, le cas bidimensionnel est expliqué en détail car il est utilisé souvent dans le manuscrit.

Dans le deuxieme chapitre, nous présentons une méthodologie pour le calcul des diagrammes de phase binaires basée sur la méthode Minima Hopping et son application pour les alliages binaires intermétalliques sodium-or et lithium-aluminium, aussi bien que pour les cristaux de chlore-hydrogène sous pressions modérées. Nous commençons par l'étude des alliages sodium-or. Les composés inter-métalliques de métaux alcalins et d'or ont des propriétés électroniques et structurelles intrigantes qui n'ont pas été largement explorées. Nous effectuons une étude systématique du diagramme de phases du système binaire Na_xAu_{1-x} appartenant à cette famille. Nous obtenons que la composition la plus stable est $NaAu_2$, en accord avec les données expérimentales. Nous confirmons également les structures cristallines de $NaAu_2$ et Na_2Au, qui ont

été entièrement caractérisées par l'expérience. Nous trouvons un candidat pour la structure de l'état fondamental de la stoechiométrie NaAu. De plus, nous obtenons trois stoechiométries, à savoir Na_3Au_2, Na_3Au et Na_5Au, qui se révèlent être thermodynamiquement stables dans nos simulations. Nous ne trouvons aucune preuve de l'existence de $NaAu_5$ qui a été discutée dans la litérature. Finalement, nous effectuons des calculs de phonons pour vérifier la stabilité de toutes les phases identifiées et des simulations des spectres de diffraction des rayons X pour permettre la comparaison avec des futures données expérimentales.

Dans la section qui suit, nous appliquons la même méthodologie aux alliages intermétalliques Li-Al. Ces derniers sont des matériaux clés pour l'ingénierie des structures légères et ont de plus été proposés comme électrodes de grande capacité pour batteries de lithium. Nous déterminons le diagramme de phase *ab initio* des cristaux de Li-Al en utilisant la méthode Minima Hopping. Au dèlá de la reproduction des phases rapportées expérimentalement (LiAl, Li_3Al_2, Li_9Al_4, $LiAl_3$, and Li_2Al), une diversité structurale plus grande que prévue est dévoilée en découvrant six phases binaires non reportées, qui sont susceptibles d'être thermodynamiquement stables. Nous avons aussi characterisé les constantes élastiques de tous les composés stables Li-Al par rapport à leur stoechiométrie.

Finalement, en étudiant les cristaux de H-Cl sous pressions moderées (100 GPa), nous trouvons que les molécules triatomiques de H, dont la stabilité est prédite pour H pure sous des pressions d'environ 2.1 TPa, deviennent stables l'intérieur de cages de Cl sous pressions beaucoup plus faibles.

Le troisième chapitre présente l'étude de matériaux avec intérêt pour l'application dans la technologie photovoltaïque: l'oxyde conducteur transparent $CuBO_2$, des nouvelles pérovskites de nitrure et les alliages de chalcopyrite $Cu(In,Ga)S_2$. Dans la première section, il est montré que le structure delafossite du $CuBO_2$ connue à ce jour, n'est pas l'état fondamental de ce matériau au contraire des conclusions très controversées de certains auteurs. Nous avons trouvé d'autres structures cristallines jusqu'à 600 meV plus basses en énergie que la delafossite.

Aprés, nous présentons nos résultats sur la possibilité de l'existence de pérovskites avec composition ABN_3. Nous trouvons 21 nouvelles compositions de la forme ABN_3 qui sont thermodynamiquement stables et qui ont donc d'excellentes chances d'être accessibles expérimentalement. La plupart des matériaux trouvés cristallisent dans des phases monocliniques, à l'exception de trois composés, notament $LaReN_3$, $LaWN_3$ et $YReN_3$ qui ont une structures du type pérovskite déformée.

Finalement, nous considérons les alliages de chalcopyrites Cu(In,Ga)S_2. La densité d'états électroniques projetée sur les orbitales s, p et d des atomes a été étudiée en fonction du rapport In/Ga, dans le cadre de la théorie de la fonctionnelle de la densité en utilisant des fonctionnelles hybrides. Des expériences de Spectroscopie de structure près du front d'absorption de rayons X effectuées sur des couches minces de $CuIn_xGa_{1-x}S_2$ pour $x = 0$, 0.67 et 1 par nos collaborateurs expérimentales, montrent un déplacement du seuil d'absorption de S avec le changement du rapport In/Ga, comme on peut s'y attendre si on considère la variation de la largeur de la bande interdite. En revanche, les seuils d'absorption de Ga, In et Cu restent indépendant de la composition. Nos calculs indiquent que ce comportement provient de la dépendance des états électroniques de l'environnement atomique local tandis que le changement dans la bande interdite surgit du changement de la moyenne spatiale de ces états localisés avec la composition de l'alliage.

Dans le chapitre 4, nous développons une fonctionnelle d'échange-corrélation semi-empirique adaptée au calcul des énergies de formation des solides dans le cadre de la théorie de la fonctionnelle de la densité. Cette fonctionnelle a la forme de celui de Perdew-Burke-Ernzerhof, mais avec trois paramètres ajustés pour reproduire les énergies de formation expérimentales d'un ensemble représentatif de composés binaires. La qualité de la fonctionnelle obtenue a été ensuite évaluée pour un ensemble de preuve. Notre fonctionnelle réussit à réduire l'erreur de Perdew-Burke-Ernzerhof par environ un facteur de deux. De plus, ce résultat est obtenu en conservant la qualité de la géométrie.

Les résultats présentés dans ce manuscrit représentent exemples significatifs du niveau de maturité atteint aujourd'hui par le domaine de la conception de nouveaux matériaux grâce à des calculs *ab initio*. La variétés de nos résultats est une preuve de la puissance et généralité de ces approches, et les développements apportés par cette thèse sont maintenant à disposition de la communauté pour la réalisation de nouvelles applications.

Table of contents

Chapter 1

Theoretical Background

Concepts and methods used along this thesis are presented here. First, we give a short introduction to the Kohn-Sham Density Functional Theory and exchange-correlation functionals. Then, in the following sections we describe different approaches to materials design focusing on the Minima Hopping Method, which was used to perform all structure prediction calculations in this work. Finally, the construction of the convex hull of stability, in particular the bidimensional case, is explained in detail as it will be used several times in the following chapters. Atomic units will be used through the chapter.

1.1 Density Functional Theory

Within the Born-Oppenheimer approximation [1], the Hamiltonian of a system of N electrons under the influence of a Coulombic field of a set of stationary nuclei can be written as (atomic units)

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}}, \tag{1.1}$$

where

$$\hat{T} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 \tag{1.2}$$

is the kinetic energy operator,

$$\hat{V}_{\text{ext}} = \sum_{i=1}^{N} v(\boldsymbol{r}_i), \tag{1.3}$$

$$v(\boldsymbol{r}_i) = -\sum_{A}^{\text{Nuclei}} \frac{Z_A}{|\boldsymbol{r}_i - \boldsymbol{R}_A|}, \tag{1.4}$$

is the sum of the potential operators of the nuclei of atomic number $\{Z_A\}$ at positions $\{R_A\}$, and

$$\hat{V}_{\text{ee}} = \sum_{i<j}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \tag{1.5}$$

is the electron-electron interaction operator.

By solving the corresponding Schrödinger equation

$$\hat{H}\Psi = E\Psi, \tag{1.6}$$

we can obtain the wavefunction Ψ and have access to all information about a system. However, the wave function is a very complicated quantity that cannot be probed experimentally and depends on three spacial and one spin variables for each particle of the system. This makes a computational treatment of the Schrödinger equation unfeasible, unless the number of particles is very small. Even just the memory to store wavefunctions would be enormous. In practice, though, usually one is not interested in all the information that the wavefuntion contains. Rather, it would be preferable to consider a simpler physical quantity to describe the system and express the expectation values of operators, making more efficient the computational treatment of the problem. It turns out that this quantity can be the electronic density which is the key ingredient of Density Functional Theory (DFT).

With that in mind, let us write the Schrödinger equation in terms of the electronic density

$$n(\boldsymbol{r}) = N \sum_{\sigma_1 ... \sigma_N} \int d\boldsymbol{r}_2 ... d\boldsymbol{r}_N |\Psi(\boldsymbol{r}\sigma_1, \boldsymbol{r}_2\sigma_2, ..., \boldsymbol{r}_N\sigma_N)|^2 \tag{1.7}$$

and the electron-pair density

$$n_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{N(N-1)}{2} \sum_{\sigma_1 ... \sigma_N} \int d\boldsymbol{r}_3 ... d\boldsymbol{r}_N |\Psi(\boldsymbol{r}_1\sigma_1, ..., \boldsymbol{r}_N\sigma_N)|^2. \tag{1.8}$$

If Ψ is the ground state of the system, supposing it is normalized, the ground state energy can be determined by multiplying on the left by Ψ^* and integrating over all positions and spin variables:

$$E = T + V_{\text{ext}} + V_{ee}, \tag{1.9}$$

where

$$T = -\frac{1}{2} \sum_{\sigma_1 ... \sigma_N} \int d\boldsymbol{r}_1 ... d\boldsymbol{r}_N \Psi^*(\boldsymbol{r}_1, ..., \boldsymbol{r}_N) \left[\sum_{i=1}^{N} \nabla_i^2 \right] \Psi(\boldsymbol{r}_1, ..., \boldsymbol{r}_N), \tag{1.10}$$

$$V_{\text{ext}} = \int d\boldsymbol{r} v(\boldsymbol{r}) n(\boldsymbol{r}), \tag{1.11}$$

and

$$V_{ee} = \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \frac{n_2(\boldsymbol{r}_1, \boldsymbol{r}_2)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \tag{1.12}$$

Equation (1.10) can be written as a function of the one-electron reduced density matrix $\gamma(\boldsymbol{r},\boldsymbol{r}') = N\sum_{\sigma_1\ldots\sigma_N}\int d\boldsymbol{r}_2 ... d\boldsymbol{r}_N \Psi^*(\boldsymbol{r}'\sigma_1,\boldsymbol{r}_2\sigma_2,...,\boldsymbol{r}_N\sigma_N)\Psi(\boldsymbol{r}\sigma_1,\boldsymbol{r}_2\sigma_2,...,\boldsymbol{r}_N\sigma_N)$ by using the antisymmetry property of the fermionic wavefunction:

$$T = -\frac{1}{2}\int d\boldsymbol{r}\left[\nabla_r^2\gamma(\boldsymbol{r},\boldsymbol{r}')\right]_{\boldsymbol{r}'=\boldsymbol{r}}. \tag{1.13}$$

From Equation (1.12) can be seen that the pair density contains all the information about exchange and correlation of the system. It is related to the probability of finding an electron at a point $\boldsymbol{r}_1$ while there is another electron at $\boldsymbol{r}_2$. If the electrons were identical non-interacting classical particles, the pair density would reduce to the product of the densities of the two electrons since the probability of finding an electron in a point would be independent of the presence of another electron. This can be used to split the density and hence the electron-electron interaction (1.12) in a classic and a quantum term

$$n_2(\boldsymbol{r}_1,\boldsymbol{r}_2) = \frac{1}{2}n(\boldsymbol{r}_1)\left[n(\boldsymbol{r}_2) + h_{\mathrm{xc}}(\boldsymbol{r}_1,\boldsymbol{r}_2)\right]. \tag{1.14}$$

The quantum term is going to contain the information of exchange and correlation while the classical part will lead to the classic electrostatic interaction between two particles. h_{xc} is called the exchange-correlation hole. Then we have

$$V_{\mathrm{ee}} = J + E^o_{\mathrm{xc}}, \tag{1.15}$$

where

$$J = \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \frac{n(\boldsymbol{r}_1)n(\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} \tag{1.16}$$

is the electrostatic repulsion of the electrons also known as Hartree energy, and

$$E_{\mathrm{xc}} = \frac{1}{2}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \frac{n(\boldsymbol{r}_1)h_{\mathrm{xc}}(\boldsymbol{r}_1,\boldsymbol{r}_2)}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} \tag{1.17}$$

is the exchange-correlation energy, which contains the quantum effects of the system.

Finally the total energy can be written in function of this quantities as as

$$E = T + V_{\mathrm{ext}} + J + E^o_{\mathrm{xc}}. \tag{1.18}$$

The formalization of the use of the density as the basic variable for the description of the system was established in a seminal work of P. Hohenberg and W. Kohn in 1964 [2] where they presented the two theorems which constitute the foundations of the DFT. The first and second Hohenberg-Kohn theorems state that (i) the external

potential $V_{\text{ext}}(\boldsymbol{r})$ is (within a constant) a unique functional of the ground state density $n(\boldsymbol{r})$; since, in turn, $V_{\text{ext}}(\boldsymbol{r})$ fixes the Hamiltonian we see that the full many particle ground state is a unique functional of $n(\boldsymbol{r})$; (ii) this density functional has its minimum at the ground state density. In other words these theorems state that the electronic density determines its own Hamiltonian, hence its own wavefunctions, and satisfies the variational principle.

Even if the Hohenberg-Kohn theorems tell us that the ground state energy is a functional of the electronic density, the explicit form of this functional remains unknown. The energy of Equation (1.18) contains the kinetic energy term that can not be written as an explicit functional of the density, and the exchange-correlation term, whose form is unknown. To address this problem, Kohn and Sham considered an auxiliary system of non-interacting electrons with the same density than the real one, and hence the same ground state energy [3]. The ground state of such a system is described by a single Slater determinant (SD) wavefunction with density

$$n(\boldsymbol{r}) = \sum_{i=1}^{N} |\phi_i(\boldsymbol{r})|^2, \tag{1.19}$$

where ϕ_i are the one-electron orbitals that form the SD, called Kohn-Sham (KS) orbitals.

If one adds and subtracts to Equation (1.18) the kinetic energy of the auxiliary system

$$T_{\text{s}} = -\frac{1}{2} \sum \langle \phi_i | \nabla^2 | \phi_i \rangle \tag{1.20}$$

and puts together all the unknown terms

$$E = T_{\text{s}} - T_{\text{s}} + T + V_{\text{ext}} + J + E_{\text{xc}}^{o}, \tag{1.21}$$

$$E = T_{\text{s}} + V_{\text{ext}} + J + \left[(T - T_{\text{s}}) + E_{\text{xc}}^{o} \right], \tag{1.22}$$

the total energy reads

$$E = T_{\text{s}} + V_{\text{ext}} + J + E_{\text{xc}}. \tag{1.23}$$

In the KS scheme the term E_{xc} contains the exchange and correlation effects plus the difference of the kinetic energy between the real and the auxiliary systems. This term is called exchange-correlation functional.

Now, in virtue of the second Hohenberg-Kohn theorem [2], the equation

$$\frac{\delta E[n(\boldsymbol{r})]}{\delta n\,(\boldsymbol{r})} = 0, \tag{1.24}$$

with the normalization of the orbitals as a constraint

$$\langle \phi_i \,|\, \phi_j \rangle = \delta_{ij}, \tag{1.25}$$

is the equation that gives the ground-state density of the interacting system and the corresponding-ground state energy. This equation can be recast in the system of equations

$$\left[-\frac{1}{2}\nabla^2 + v(\boldsymbol{r}) + \int d\boldsymbol{r}_1 \frac{n(\boldsymbol{r}_1)}{|\boldsymbol{r} - \boldsymbol{r}_1|} + v_{\mathrm{xc}}(\boldsymbol{r}) \right] \phi_i(\boldsymbol{r}) = \epsilon_i \phi_i(\boldsymbol{r}), \tag{1.26}$$

known as the KS equations, where

$$v_{\mathrm{xc}}(\boldsymbol{r}) = \frac{\delta E_{\mathrm{xc}}[n(\boldsymbol{r})]}{\delta n\,(\boldsymbol{r})}. \tag{1.27}$$

The KS equations can be seen as the Schrödinger equation of a system of non-interacting electrons under the external potential

$$v_{\mathrm{KS}}(\boldsymbol{r}) = v(\boldsymbol{r}) + \int d\boldsymbol{r}_1 \frac{n(\boldsymbol{r}_1)}{|\boldsymbol{r} - \boldsymbol{r}_1|} + v_{\mathrm{xc}}(\boldsymbol{r}) \tag{1.28}$$

and its eigenvalues ϵ_i are the orbital energy levels. The solution has to be achieved self-consistently since its potential depends on the electronic density which is calculated from the single-particle wavefunction as in (1.19). The KS theory is exact. Up to this point we have not made any approximation. However, the exchange-correlation functional and hence v_{xc}, are unknown. The search for improved exchange-correlation functionals remains as a challenge in DFT.

To solve the KS equations in the case of periodic systems, it is convenient to require that the single-particle wavefunctions obey periodic boundary conditions. In virtue of the Bloch theorem, the KS orbitals have the form

$$\phi_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n\boldsymbol{k}}(\boldsymbol{r}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{1.29}$$

where $u_{n\boldsymbol{k}}(\boldsymbol{r})$ is a periodic function with the same periodicity of the lattice and $\boldsymbol{k}$ is the crystal wavevector and lies inside the first Brillouin zone (BZ).

Since any periodic function can be expanded in Fourier series, $u_k(\boldsymbol{r})$ can be written as

$$u_{nk}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} e^{i\boldsymbol{G}\cdot\boldsymbol{r}}, \tag{1.30}$$

where the sum is over the integer multiples $\boldsymbol{G} = m_1\boldsymbol{G}_1 + m_2\boldsymbol{G}_2 + m_3\boldsymbol{G}_3$ of the reciprocal lattice vectors

$$\boldsymbol{G}_1 = 2\pi \frac{\boldsymbol{R}_2 \times \boldsymbol{R}_3}{\boldsymbol{R}_1 \cdot (\boldsymbol{R}_2 \times \boldsymbol{R}_3)},\ \boldsymbol{G}_2 = 2\pi \frac{\boldsymbol{R}_3 \times \boldsymbol{R}_1}{\boldsymbol{R}_2 \cdot (\boldsymbol{R}_3 \times \boldsymbol{R}_1)},\ \boldsymbol{G}_3 = 2\pi \frac{\boldsymbol{R}_1 \times \boldsymbol{R}_2}{\boldsymbol{R}_3 \cdot (\boldsymbol{R}_1 \times \boldsymbol{R}_2)},$$

and $\boldsymbol{R}_1$, $\boldsymbol{R}_2$ and $\boldsymbol{R}_3$ are the lattice vectors. Substituting (1.30) in (1.29), the KS orbitals can be written as

$$\phi_{nk}(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}}. \tag{1.31}$$

To write the KS equation in a plane wave basis, we first consider the effect of the Laplacian operator on the Fourier expansion of the orbitals

$$-\frac{1}{2}\nabla^2 \left[\sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}} \right] = \frac{1}{2} \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} \left[\boldsymbol{G}+\boldsymbol{k}\right]^2 e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}} \tag{1.32}$$

and insert (1.32) in the KS equation

$$\frac{1}{2} \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} \left[\boldsymbol{G}+\boldsymbol{k}\right]^2 e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}} + v_{\mathrm{KS}}(\boldsymbol{r})[n(\boldsymbol{r})] \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}} = \epsilon_{nk} \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} e^{i(\boldsymbol{G}+\boldsymbol{k})\cdot\boldsymbol{r}}.$$

Multiplying on the left by $e^{-i(\boldsymbol{G}'+\boldsymbol{k})\cdot\boldsymbol{r}}$ and integrating over $\boldsymbol{r}$ we obtain

$$\frac{1}{2} \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}^{n\boldsymbol{k}} \left[\boldsymbol{G}+\boldsymbol{k}\right]^2 \delta_{\boldsymbol{G}\boldsymbol{G}'} + \sum_{\boldsymbol{G}} v_{\mathrm{KS}}(\boldsymbol{G},\boldsymbol{G}') C_{\boldsymbol{G}}^{n\boldsymbol{k}} = \epsilon_{nk} \delta_{\boldsymbol{G}\boldsymbol{G}'} C_{\boldsymbol{G}}^{n\boldsymbol{k}},$$

where

$$v_{\mathrm{KS}}(\boldsymbol{G},\boldsymbol{G}') = \int d\boldsymbol{r} e^{-i\boldsymbol{G}'\cdot\boldsymbol{r}} v_{\mathrm{KS}}(\boldsymbol{r})[n(\boldsymbol{r})] e^{i\boldsymbol{G}\cdot\boldsymbol{r}} \tag{1.33}$$

is the Fourier transformation of the KS potential and $[n(\boldsymbol{r})]$ represents the functional dependence of v_{KS} on the electronic density.

Finally, the KS equation for a periodic system in a plane-wave basis can be written as

$$\frac{1}{2} \left[\boldsymbol{G}+\boldsymbol{k}\right]^2 C_{\boldsymbol{G}}^{n\boldsymbol{k}} + \sum_{\boldsymbol{G}} v_{\mathrm{KS}}(\boldsymbol{G},\boldsymbol{G}') C_{\boldsymbol{G}}^{n\boldsymbol{k}} = \epsilon_{nk} C_{\boldsymbol{G}}^{n\boldsymbol{k}}. \tag{1.34}$$

Equation (1.34) has to be solved self-consistently to find the $C_{\boldsymbol{G}}^{n\boldsymbol{k}}$ coefficients for each $\boldsymbol{k}$ vector.

The equation is generally solved in a discretized reciprocal space for each $\boldsymbol{k}$ vector in the first BZ (or for the symmetrically inequivalent ones, after considering the symmetries of the crystal). Also the plane wave expansion is truncated at a certain $\boldsymbol{G}$ vector: the plane waves with large kinetic energy are neglected in the Fourier expansion. From Equation (1.32), it can be seen that the terms with the larger kinetic energy are the ones corresponding to large $\boldsymbol{G}$ vector.

1.2 Exchange-correlation Functionals

The exchange-correlation functional E_{xc} of Equation (1.17) is usually separated into an exchange and a correlation part [4, 5]

$$E_{\mathrm{xc}}[n] = E_{\mathrm{x}}[n] + E_{\mathrm{c}}[n], \tag{1.35}$$

where the exchange is defined as [4, 6]

$$E_{\mathrm{x}}[n] = \langle \Phi^{\min} | \, \hat{V}_{\mathrm{ee}} \, | \Phi^{\min} \rangle - J[n] \tag{1.36}$$

and $\Phi^{\min}$ is the KS wavefunction. The correlation part is

$$E_{\mathrm{c}}[n] = E_{\mathrm{xc}}[n] - E_{\mathrm{x}}[n] = \langle \Psi^{\min} | \, \hat{V}_{\mathrm{ee}} \, | \Psi^{\min} \rangle - \langle \Phi^{\min} | \, \hat{V}_{\mathrm{ee}} \, | \Phi^{\min} \rangle \,, \tag{1.37}$$

where $\Psi^{\min}$ is the exact wavefunction of the interacting system.

Regardless E_{x} is known exactly, it is not an explicit functional of the density and the derivative (1.27) can not be readily evaluated. For this reason, we need to approximate both E_{x} and E_{c}. Furthermore, in practice it is difficult to achieve the desired accuracy by combining exact exchange with approximate correlation.

The Local Density Approximation

The simplest approach to exchange-correlation functionals is the Local Density Approximation (LDA) [2]. It was proposed by Kohn and Sham [2], and proved to be remarkably accurate, useful and hard to improve upon. It consists in applying the exact results for the uniform electron gas to real systems.

The LDA functional has the form

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int d\boldsymbol{r} n(\boldsymbol{r}) \epsilon_{\mathrm{xc}}[n(\boldsymbol{r})], \tag{1.38}$$

where $\epsilon_{\mathrm{xc}}[n(\boldsymbol{r})] = \epsilon_{\mathrm{x}}[n(\boldsymbol{r})] + \epsilon_{\mathrm{c}}[n(\boldsymbol{r})]$ is the exchange-correlation energy density per particle of the uniform electron gas evaluated at the density of the system at the point $\boldsymbol{r}$. As the LDA is exact for a uniform density, it is expected to be particularly accurate for a density that varies slowly in space.

The LDA exchange energy per particle is

$$\epsilon_{\mathrm{x}}[n] = -\frac{3}{4}\left(\frac{3}{2\pi}\right)^{2/3} \frac{1}{r_s}, \tag{1.39}$$

where $r_s = (3/4\pi n)^{1/3}$ is the radius of a sphere which in average contains one electron, known ad the Wigner-Seitz radius.

The extension to spin-polarized system is called the Local Spin Density Approximation (LSDA). In this case the exchange energy is cast in the form

$$E_{\mathrm{x}}^{\mathrm{LSDA}}[n_\uparrow, n_\downarrow] = \int d\boldsymbol{r} n(\boldsymbol{r}) \epsilon_{\mathrm{x}}[\, n, \zeta\,], \tag{1.40}$$

where

$$\epsilon_{\mathrm{x}}[\, n, \zeta\,] = -\frac{3}{4}\left(\frac{3}{2\pi}\right)^{2/3} \frac{1}{r_s} \frac{\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right]}{2}, \tag{1.41}$$

$n = n_\uparrow + n_\downarrow$ and $\zeta = [n_\uparrow - n_\downarrow]/n$. For the spin-compensated case ($\zeta = 0$), the expression (1.39) is obtained, and for the spin-fully-polarized case ($\zeta = 1$)

$$\epsilon_{\mathrm{x}}[n] = -\frac{3}{4}\left(\frac{3}{\sqrt{2}\pi}\right)^{2/3} \frac{1}{r_s}. \tag{1.42}$$

For intermediate spin-polarizations ($0 < \zeta < 1$) , ϵ_x can be written as an interpolation between those two cases

$$\epsilon_{\mathrm{x}}[\, n, \zeta\,] = \epsilon_{\mathrm{x}}(\, n, \zeta = 0\,) + [\epsilon_{\mathrm{x}}(\, n, \zeta = 1\,) - \epsilon_{\mathrm{x}}(\, n, \zeta = 0\,)]\, f[\zeta], \tag{1.43}$$

where

$$f(\zeta) = \frac{1}{2}\frac{(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2}{2^{1/3} - 1}. \tag{1.44}$$

Correlation is a more difficult problem. The correlation energy per particle $\epsilon_{\mathrm{c}}[n]$ is known only for two limiting cases: the high-density (weak-coupling) and the low-density

(strong-coupling) limits of a spin-compensated uniform electron gas. The expressions are

$$\epsilon_{\rm c}[r_s \ll 1] = c_0 \ln(r_s) + c_1 + c_2 r_s \ln(r_s) + c_3 r_s + \cdots \tag{1.45}$$

and

$$\epsilon_{\rm c}[r_s \gg 1] = \frac{1}{2}\left(\frac{d_0}{r_s} + \frac{d_1}{r_s^{3/2}} + \frac{d_2}{r_s^2} + \cdots\right). \tag{1.46}$$

The exact numerical values of $\epsilon_{\rm c}[r_s \ll 1]$ and $\epsilon_{\rm c}[r_s \gg 1]$ are known with small statistical uncertainties for several intermediate values of r_s from Monte Carlo simulations of the uniform electron gas carried out by Ceperley and Alder [7]. Based on these results, several interpolation formulas have been devised to connect the high- and low-density limits.

The most used analytic expression satisfying both limits was proposed by Perdew and Wang by fitting the Ceperley-Alder data for r_s=2, 5, 10, 20 and 100 [8]:

$$\epsilon_{\rm c}(n) = -2c_0(1+\alpha_1 r_s)\ln\left[1 + \frac{1}{2c_0(\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_1 r_s^2)}\right], \tag{1.47}$$

where $c_0 = 0.031091$, $\alpha_1 = 0.21370$, $\beta_3 = 1.6382$, $\beta_4 = 0.49294$, $\beta_1 = 1/2c_0\exp(-c_1/2c_0)$, $\beta_2 = 2c_0\beta_1^2$, and $c_1 = 0.0466644$.

Analogously to Equation (1.40), the extension of the correlation energy to spin-polarized system can be written as

$$E_{\rm c}^{\rm LSDA}[n_\uparrow, n_\downarrow] = \int d\boldsymbol{r} n(\boldsymbol{r}) \epsilon_{\rm c}(r_s, \zeta). \tag{1.48}$$

However, unlike for exchange, there is no simple exact formula relating $\epsilon_c(r_s, \zeta = 0)$ and $\epsilon_{\rm c}(r_s, \zeta = 1)$. Vosko, Wilk and Nusair [9] recommended the following expression

$$E_{\rm c}^{\rm LSDA}(r_s, \zeta) = \epsilon_{\rm c}(r_s, \zeta = 0) + \alpha_c[r_s]\frac{f(\zeta)}{f''(0)}(1-\zeta^4) + [\epsilon_{\rm c}(r_s, \zeta = 1) - \epsilon_{\rm c}(r_s, \zeta = 0)]\, f(\zeta)\zeta^4, \tag{1.49}$$

where $\alpha_c[r_s] = \partial^2 \epsilon_{\rm c}(r_s, \zeta)/\partial\zeta^2$ is called spin stiffness.

The LSDA is exact for a uniform electron gas and quite accurate for solids, but less satisfactory for atoms and molecules.

Generalized Gradient Approximation

The next natural step in the improvement of the approximate expressions for the exchange-correlation functionals is to expand $E_{\rm xc}$ in gradients for the density. $E_{\rm x}$ and

$E_{\rm c}$ can be written as

$$E_{\rm x}^{\rm GE}(n) = \int d\boldsymbol{r} n \epsilon_{\rm x}^{\rm LDA}(n) \left[1 + \mu s^2 + \cdots\right], \tag{1.50}$$

$$E_{\rm c}^{\rm GE}(n) = \int d\boldsymbol{r} n \left[\epsilon_{\rm c}^{\rm LDA}(n) + \beta(n) t^2 + \cdots\right], \tag{1.51}$$

where $s = |\nabla n|/(2(3\pi^2)^{1/3} n^{4/3})$ and $t = |\nabla n|/4(3/\pi)^{1/6} n^{7/6}$ are reduced density gradients [4, 6].

However, the density-gradient expansion results only in modest improvements over LDA for exchange energies [4]. The second-order density gradient expansion for exchange performs well only in the limit of small reduced density gradients s. The assumption that the reduced gradients are small may be justified for an infinite electron gas but not for finite systems. Besides, several spurious features appear: the exchange hole violates the negativity constraint and it does not integrate to -1 [10]. The correction needed to reproduce the exact correlation energy is overestimated by a factor of 5 leading to the prediction of positive correlation energies [4, 6]. Because of this, truncated density-gradient expansions are not used as practical density functionals. Instead, they are regarded as the exact forms to which approximate exchange-correlation functionals should reduce in the limit of slowly-varying densities [4, 6].

A different and successful strategy is to look for gradient corrected functionals of the form [10]

$$E_{\rm xc}^{\rm GGA}[n_\uparrow, n_\downarrow] = \int d\boldsymbol{r} f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow), \tag{1.52}$$

where the function f is constructed to satisfy important exact constraints like asymptotic behavior and boundaries of $\epsilon_{\rm xc}$, density scaling transformations or to fit properties of a particular class of systems. Such functionals are called Generalized Gradient Approximations (GGA)[10, 11]. In comparison to LSDA, GGA functionals tend to improve total energies, atomization energies, energy barriers and structural energy differences.

The GGA for exchange and for correlation are written as

$$E_{\rm x}^{\rm GGA}[n] = \int d\boldsymbol{r} n \epsilon_{\rm x}^{\rm unif}[n] F_{\rm x}(s) \tag{1.53}$$

and

$$E_{\rm xc}^{\rm GGA}[n_\uparrow, n_\downarrow] = \int d\boldsymbol{r} n [\epsilon_{\rm c}^{\rm unif}(r_s, \zeta) + H(r_s, \zeta, t)]. \tag{1.54}$$

respectively, where F_{x} is called enhancement factor and $H(r_s, \zeta, t)$ is the gradient contribution to correlation.

Many analytic properties of the exact exchange-correlation functional are known and, as mentioned before, they can be used to determine F_{x} and H (the function $f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow)$) [4, 6, 12]. For example:
(i) the exchange energy is strictly negative, while the correlation energy is non-positive

$$E_{\mathrm{x}} < 0 \qquad E_{\mathrm{c}} \leq 0, \tag{1.55}$$

(ii) the Lieb-Oxford bound [13]

$$E_{\mathrm{x}}[n_\uparrow, n_\downarrow] \geq E_{\mathrm{xc}}[n_\uparrow, n_\downarrow] \geq 1.679 \int d\boldsymbol{r} n(\boldsymbol{r})^{4/3}, \tag{1.56}$$

(iii) in a one-electron system of density n_1, E_{x} must cancel the Coulomb self-interaction energy, while E_{c} must vanish

$$E_{\mathrm{xc}}[n_1] + J[n_1] = 0 \qquad E_{\mathrm{c}}[n_1] = 0, \tag{1.57}$$

(iv) GGA functional expressions must recover the LSDA ones for the uniform electron gas.

Many properties of density functionals are derived from density scaling transformations. The uniform scaling of the density is defined as

$$n_\gamma(\boldsymbol{r}) = \gamma^3 n(\gamma \boldsymbol{r}), \tag{1.58}$$

as it integrates also to the number of electrons. Considering the effect of such transformations in the exchange-correlation functional leads to important constraints like [4, 6]

$$E_{\mathrm{x}}[n_\gamma] = \gamma E_{\mathrm{x}}[n] \qquad \lim_{\gamma \to \infty} E_{\mathrm{c}}[n_\gamma] > -\infty. \tag{1.59}$$

In Chapter 3, we will discuss the constraint satisfaction procedure in detail for the derivation of the PBE functional [12]

1.3 First-Principles Cristal Structure Prediction

Crystal structure prediction is the problem of determining the ground state configuration of a given system of atoms in a unit cell, i.e. to determine the position of the atoms and the lattice parameters. This is an optimization problem, where the function

to minimize is the potential energy and the variables are the atomic positions and the cell parameters [14, 15].

Even nowadays, when first-principle modeling of materials has become routine, the prediction of a new crystal structure continues being a challenge, which requires highly efficient algorithms capable of finding a way from a trial initial structure towards the ground state geometry. Of course, using a traditional optimization algorithm like steepest descent or conjugate gradients to minimize the energy would only take the system from the initial configuration to the nearest energy minimum, which is extremely unlikely to be the global one. One solution is to try different starting geometries, but reaching the global minimum in this way certainly needs a smart way of selecting the trial structures.

A possible approach is to turn to chemical principles and empirical rules. Several chemistry constraints can be imposed to screen among the candidates. Many crystals share the same structure and motifs, therefore it is reasonable to choose known structures as initial guess. This idea is used in high throughput structure prediction where new stoichiometries are tested in large databases of known structures often referred as prototypes [16, 17].

Likewise, candidate structures can be randomly generated but by satisfying chemical constraints. Setting a range of unit cell volumes according to the number of atoms and a range of distances between the atoms, are strong constraints to help to limit the infinite number of random structures. It is also common to enforce space groups and building blocks to bias the search. This is known as the Random Search (RS) method [15, 18]. In general, RS is performed combined with the use of prototypes [19].

More advanced techniques go beyond chemistry rules and use evolutionary principles such as inheritance, mutation, selection, and crossover to generate new candidate structures. This is the case of genetic algorithms. They typically start with a set of (parent) structures from which new ones are going to be generated (children) by modifying and combining them in such a way that they minimize an objective function, in the case of structure prediction, the energy. For instance, a weighted average of the lattice parameter of the parents can be used to determine the ones of the children. Children can be also formed by deformations of a single parent, as permutation of atoms of different kind, modification of the unit cell, etc. Then the lowest-energy structures of a generation are chosen to procreate the next generation. Analogously to evolution in life, the bad features will diminish and the candidate structures will evolve towards the global minimum of the potential energy.

Particle swarm optimization [20] has also been used for structure prediction. This algorithm shares some similarities with evolutionary techniques, it is initialized with a population of random structures and searches for minima by updating generations. However, this algorithm has no evolution operators such as crossover and mutation [21, 22].

Other methods take a rather different approach to explore the potential energy surface (PES). This is the case of algorithms like Simulated Annealing (SA)[23] and Basin Hopping (BH) [24]. In those algorithms the PES is explored by overcoming energy barriers. SA starts from a random structure, the atoms are randomly displaced and the unit cell parameters are changed. After a perturbation (atoms positions or cell parameters), the energy of the new resulting configuration is calculated. Then, the new configuration will be accepted or rejected if the probability $P = e^{-\frac{\Delta E}{k_B T}}$ is larger or lower than a random number between 0 and 1, ΔE is the difference between the energies of the new and the current configurations, k_B is the Boltzmann constant and T is the simulation temperature. Accepting a configuration means that the next perturbation will be done from this configuration, and rejecting it, means that the perturbation is neglected, and another perturbation will be performed starting from the current configuration. In other variants of SA, the PES is explored with molecular dynamics instead of the Monte Carlo method described above. The simulation starts with a high temperature in such a way that most of the configurations will be accepted. This allows to overcome high energies barriers. Then, the temperature is decreased gradually during the simulation, and higher energies configurations are less accepted. Finally, when the temperature reaches zero, only configurations that lower the energy are accepted and the system ends up in the closest minimum. In practice, SA runs tend to be very slow. Moreover, the number of minima increases exponentially with the number of atoms. The BH algorithm explores the PES with a Monte Carlo method as in SA, but after each perturbation the geometry is relaxed, i.e. the system is taken to the minimum of the basin it has arrived, and the energy of the relaxed structures are used to compute the probability of acceptance. The temperature is maintained constant during the calculation, which poses the problem of finding its optimal value for the run. In BH there is no penalty to avoid revisiting basins.

Another efficient method of structure prediction is Minima Hopping, where a molecular dynamic algorithm is combined with chemical principles to optimize the search for the global minimum of the potential energy. In the next section, Minima Hopping is discussed in detail as it is the method used in this thesis to perform all the calculations of structure prediction.

1.4 The Minima Hopping Method

Minima Hopping (MH) [25, 26] is a global optimization algorithm based on the principle of exploring the high-dimensional PES of complex systems of atoms as fast as possible and avoiding revisiting minima while progressing towards the (global) minimum structure. It is constituted by an inner part that performs jumps between the minima to escape from the basins and an outer part that rejects or accepts the new minima according to a threshold-based rule, such that half of the moves are accepted and half are rejected, which is achieved by consequently modifying the threshold parameter. A record of the visited minima is kept.

The jumps consist of a short Molecular Dynamics (MD) run followed by a geometry relaxation performed by a combination of standard steepest descent and conjugate gradient. The MD part is used to move the system from one basin of the PES to another by crossing energy barriers. MD needs an initial velocity distribution, which is then scaled to fit the desired kinetic energy. In the first implementation of MH, the velocities were randomly directed for each atom with Gaussian distributed magnitudes, but in successive versions the direction of the velocities started to be biased using a method called "softening", which will be described later in this section. Then, after a MD escape, the local geometry relaxation takes the system to the minimum of the basin it has arrived. This may lead to three cases. The first one is that it may give back the local minimum used as starting point. The second case is that the new minimum is one that was previously visited. The third case is when the minimum is a new one, i.e. it has not been previously visited. This last case is the desirable one since it will result in the exploration of new configurations.

Once a new minimum is reached, the outer part is going to either accept it or reject it. If the new minimum is lower than the starting one by a certain value E_d, it is going to be accepted, this means that the system will move to that minimum from which another MD escape will start. Otherwise the system will not move to it, the kinetic energy will be increased and then another MD escape will start from the same minimum. The outer part introduces a preference for moves that go down in energy. However, if the inner part only proposes moves that go up in energy, the outer part will eventually accept them, since E_d is going to be increased to compensate for that many rejections. The kinetic energy is not only increased if the system falls back into the current minimum, but also if it gets into another minimum that was already visited.

It might seem reasonable to use a very large kinetic energy since it is desirable to explore new configurations, but this is only true if the height of a new local minimum behind a barrier was independent of the barrier height. In chemistry, the Bell-Evans-

Polanyi (BEP) principle states that strongly exothermic reactions have a low activation energy. Reactants and products are neighboring local minima on the PES separated by an energy barrier. Generalizing the BEP principle to any transitions between local minima on the PES during MD simulations, it has been found that by crossing over high barriers it is more unlikely to end up in a low energy minimum [27]. MH takes advantage of that to reduce the space to be explored before finding the global minimum. For this reason it is important to make a compromise in the choice of the kinetic energy to obtain the shortest possible simulation time. If it is set to a very large value, the space to explore before arriving to the global minimum may become very large. In practice, it turns out that the best solution is to adjust dynamically the kinetic energy in such a way that half of the MD simulations lead to new basins.

Furthermore, it has been seen that low energy barriers are generally connected to low frequency eigenmodes of local minima [28]. This take us back to the assignation of the velocities before a MD simulation. To take advantage of that, before starting a MD escape, MH performs a search for the lowest curvature direction to escape from the current minimum. This is the method “softening" [26, 29] that we mentioned before. The softening is efficiently done by an iterative dimer method which does not need second derivatives, it uses just gradients instead. Softening has proved to significantly reduce the number of steps to find the global minimum with respect to the use of randomly-oriented velocities. However, if the softening process is executed until it converges exactly to the lowest-curvature direction, the performance drops again. Hence, after a few steps the iterations have to be stopped before the optimal lowest local curvature mode is found.

The MH algorithm has proved to be very efficient, being able to reproduce experimental results and to find novel structures starting only from the chemical composition as the only input information [30–36].

Figure 1.1 shows the evolution of the minima in a MH calculation to find the ground state structure of LiAl. For this compound the ground state is experimentally known and we can see at which step the global minimum was found. In the case of new materials, there is no experimental information to compare with and it is unknown whether the lowest energy structure found by the algorithm is actually the global minimum. However, in many cases there are indications that suggest that the global minimum was found as we discuss bellow.

In the case of LiAl, the initial geometry was selected as the lowest-volume structure from a large set of randomly-generated cells with 2 atoms of Li and 2 atoms of Al, with the constraint that the atoms are not too close (interatomic distances at least equal

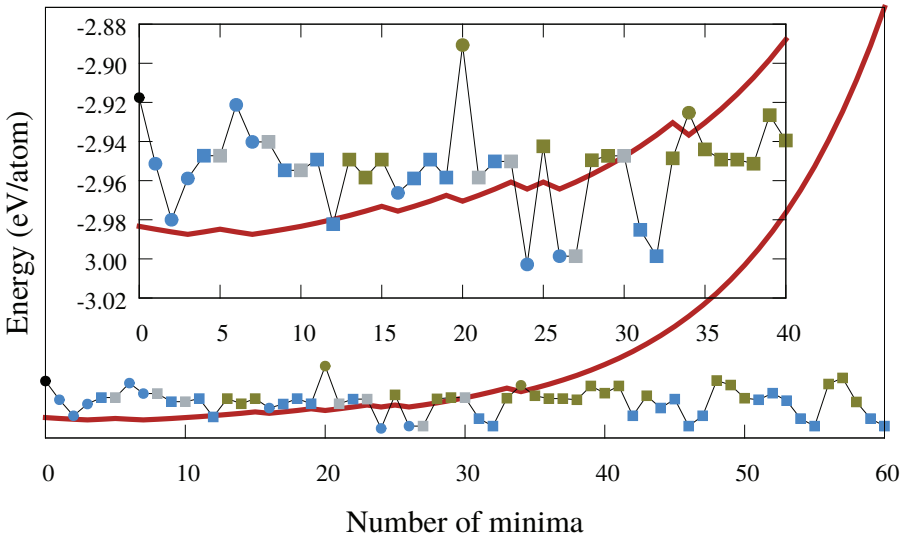


Fig. 1.1 Energy of the minima ordered by time occurrence in a MH calculation of 2 formula units of LiAl. The inset is a zoom of the first 40 minima. Accepted minima are represented by the blue symbols and rejected minima, with the green ones. The gray symbols correspond to the steps where the MD scape was not able to jump out of the basin and the same minimum occurs again. The first time a minima is visited is represented with a circle, while the revisited ones are squares. The dark red line represents the evolution of the kinetic energy (in arbitrary units).

to the sum of their covalent radii). The structure was relaxed obtaining the energy represented with the black dot in Figure 1.1. The energies and forces were calculated using the projector augmented wave (PAW) method as implemented in the code VASP [37, 38] with the Perdew-Burke-Ernzerhof (PBE)[12] approximation to the exchange correlation functional.

The ground state structure was the 24^{th} minimum to be found. This search was very fast as LiAl is a small and simple system (4 atoms per unit cell with PAW datasets for both for Li and Al with 3 electrons), for larger systems, in general, it is necessary to visit many more minima to find the global one. It can be seen in Figure 1.1 how the temperature starts increasing monotonically after the step 34, when most of the minima were already visited (square symbols predominate). This is an indicator that the global minimum might have been found. However, on MH calculations, the global minimum is searched for a fixed number of atoms and it may happen that the true ground state corresponds to a unit cell with larger number of atoms. As we are limited by computational cost, we run several unit cells with up to 10 or 12 atoms for each stoichiometry. That is Li_2Al_2, Li_3Al_3, Li_4Al_4, Li_5Al_5 and Li_6Al_6. We normally do not

run the 1-1 case as the 2-2 is still very fast. In this case all yielded the same ground state structure.

In the inset of Figure 1.1, it can be seen in detail how the temperature is decreased after jumping to lower minima, and increased after jumping to higher minima or when the MD step is not able to take the system out of a basin. Until the 12^{th} minimum, most of the minima are accepted, even when jumps were from a lower minima as from the 2^{th} to the 3^{th} minima.

The MH algorithm can be summarized as follows

```
Initialize a random cell with the atoms at random
positions and with velocities from a Boltzmann distribution

MDstart
ESCAPE TRIAL PART

  Softening

  Start a MD trajectory with kinetic energy Ekinetic from the current minimum Mcurrent.
  Once the potential energy reaches the mdmin-th minimum along the trajectory, stop MD
  and optimize geometry to find the closest local minimum 'M'

  if (M == Mcurrent) then
    Ekinetic = Ekinetic*beta1  (beta1 > 1)
    go to MDstart
  else if (M == minimum_visited_previously) then
    Ekinetic = Ekinetic*beta2  (beta2 > 1)
    go to MDstart
  else if (M == new_minimum) then
    Ekinetic = Ekinetic*beta3  (beta3 < 1)
  end if

DOWNWARD PREFERENCE PART
  if (energy[M]-energy[Mcurrent] < Ed) then
    accept new_minimum: Mcurrent=M
    add Mcurrent to history list
    Ediff=Ed*alpha1 (alpha1 < 1)
  else
    Ediff=Ediff*alpha2 (alpha2 > 1)
 end if
```

1.5 The Convex Hull of Phase Stability

The standard enthalpy of formation or heat of formation of a compound is the change of enthalpy during the formation of 1 mole of the compound from its constituent elements, with all substances in their standard states at 101.3 kPa (usually at 298.15 K). It is the central magnitude on the *ab initio* study of stability since it tells how favorable is the formation of a phase with respect to the parts from which it is composed.

Let us first briefly describe how it is determined experimentally. The enthalpy of formation can be measured by calorimetry experiments. The reactants are added to a calorimeter and the reaction is started. The calorimeter is a container with good insulated walls to prevent heat exchange with the environment and its specific heat has been determined experimentally. Usually the specific heat of the calorimeter is expressed in equivalent of moles of water. Then, measuring the difference in temperature before and after the reaction the heat of formation can be calculated as

$$\Delta H_f^o = n_W C_{pW}^o \Delta T,$$

were n_W is the number of moles of water that would absorb the same heat as the calorimeter and C_{pW}^o is the specific heat of water.

It is also possible to determine the enthalpy of formation of compounds involved in chemical reactions with other compounds whose enthalpy of formation has been already measured

$$\Delta H_f^o = \sum n_p \Delta H_p^o - \sum n_r \Delta H_r^o,$$

where n is the amount of substance, and r and p stand for reactant and products respectively. The enthalpy of formation of elementary substances is equal to zero by definition.

In the DFT framework, the enthalpy of formation of condensed systems is determined from the ground state energy at zero temperature plus the pressure times the cell volume. If no pressure is applied, usually one talks about formation energy, since the term PV, will be zero and the enthalpy will be equal to the energy.

Just by looking at the formation energy values, one has not enough information to know whether a certain phase is stable or not. The formation energy measures the stability of a phase with respect to the elements which composes it. To determine if a phase is stable, one has to consider also the decomposition into all possible non-elementary phases. In the case of binary systems, a 2D convex hull diagram is a convenient way to visualize the stability between the phases.

The general definition of the convex hull of a set of points S in n dimensions is the intersection of all convex sets containing S. In this thesis, as it is used to display the phase diagram of binary systems, we only need the bidimensional case which can be defined as the smallest convex polygon that contains all the points of S that lie in a plane. A thermodynamic stability diagram of a binary compound is nothing but the convex hull formed by the lines connecting the lowest energy points in the plot of the

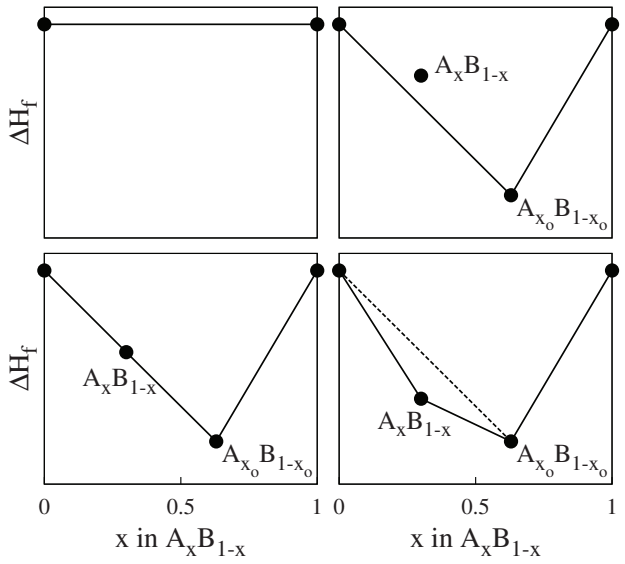


Fig. 1.2 Example of convex hull construction for a binary system AB. (top left) there are no stable phases, (top right) the phase of concentration x is unstable and decomposes in the phase at x_o plus precipitation of B, (bottom left) A_xB_{1-x} is in equilibrium with $A_{x_o}B_{1-x_o}$, i.e. the system does not loose or gain energy with this phase transformation. (bottom,right) A_xB_{1-x} and $A_{x_o}B_{1-x_o}$ are stable.

formation energy per atom as a function of x in A_xB_{1-x} (Figure 1.2):

$$\Delta H_f\left(A_xB_{1-x}\right) = E\left(A_xB_{1-x}\right) - \left[xE^o\left(A\right) + \left(1-x\right)E^o\left(B\right)\right],$$

with x in the interval [0;1], E^o the energy per atom, E the total energy and ΔH_f is the formation energy per atom. ΔH_f has a linear dependence on x.

To understand how the convex hull is helpful for the determination of the relative stability of the phases, let us examine Figure 1.2. Consider the points corresponding to the pure elements, x=0 and x=1. For x=0 we have

$$\Delta H_f\left(x\right) = E\left(B\right) - \left[0E^o\left(A\right) + E^o\left(B\right)\right] = 0 \tag{1.60}$$

and for x=1 we have

$$\Delta H_f\left(x\right) = E\left(A\right) - \left[E^o\left(A\right) + 0E^o\left(B\right)\right] = 0. \tag{1.61}$$

The line ΔH_f=0 connecting the extreme points of ΔH_f as a function of x,

$$E\left(A_xB_{1-x}\right) = xE^o\left(A\right) + \left(1-x\right)E^o\left(B\right),$$

corresponds to the set of phases that are in equilibrium with the separated elements. Their energy of formation is 0 since there is no need to give energy to the system to obtain a separation of phases. The points above this lines correspond to unstable phases.

In contrast, stable phases correspond to points under the line, i.e. negative energies of formation. If the outcome of a structure prediction calculation, like MH, are two structures of the same concentration x, it is easy to distinguish which one is more stable: the one with lower energy of formation. But when it comes to compare structures with different concentration, it is not trivial, and it becomes clear the importance of the convex hull construction.

Consider two phases with different negative values of energies of formation and concentrations. The possible dissociation of the phase with higher energy of formation (less negative) is described by the following chemical equation:

$$A_x B_{1-x} \rightarrow \left(1 - \frac{x}{x_o}\right) B + \frac{x}{x_o} A_{x_o} B_{1-x_o}. \tag{1.62}$$

Depending on the energies of each term on the equation, there will occur one of the possibilities considered before:

$$E(A_x B_{1-x}) \begin{cases} < \left(1 - \frac{x}{x_o}\right) E(B) + \frac{x}{x_o} E(A_{x_o} B_{1-x_o}) & A_x B_{1-x} \text{ is stable,} \\ = \left(1 - \frac{x}{x_o}\right) E(B) + \frac{x}{x_o} E(A_{x_o} B_{1-x_o}) & \text{both phases are in equilibrium,} \\ > \left(1 - \frac{x}{x_o}\right) E(B) + \frac{x}{x_o} E(A_{x_o} B_{1-x_o}) & A_x B_{1-x} \text{ is unstable.} \end{cases}$$

The relation of this expression with the convex hull can be easily understood by subtracting $xE(A) + (1-x)\,E(B)$ on each side:

$$\begin{aligned} E(A_x B_{1-x}) - xE(A) - (1-x)E(B) = \left(1 - \frac{x}{x_o}\right) E(B) + \frac{x}{x_o} E(A_{x_o} B_{1-x_o}) - \\ - xE(A) - (1-x)E(B) \end{aligned}$$

$$\begin{aligned} = \frac{x}{x_o} [E(A_{x_o} B_{1-x_o}) - x_o E(A) \\ - (1-x_o)E(B)]. \end{aligned}$$

In each side we have the formation of each phase respectively

$$\Delta H_f(A_x B_{1-x}) = \frac{x}{x_o}\Delta H_f(A_{x_o} B_{1-x_o}). \tag{1.63}$$

This expression is the equation of the line that connects the point at x with the one at x_o in the convex hull. This translates the cases considered before to a geometrical description: the phase at x is going to be stable if its formation energy lies in/or under the line that connects the stable phase at x_o with the point at 0 concentration of A. If it is above the line, it is going to be unstable.

Now it is clear to see that Figure 1.2 shows the cases examined before: when there are no stable phases, when the phase of concentration x is unstable and decomposes in the phase at x_o plus precipitation of B, when both phases are stable or both phases are in equilibrium, i.e. the system does not loose or gain energy with this phase transformation.

Analogously, if the phase at concentration x were at the right side of the one at x_o ($x > x_o$), the chemical equation would be

$$A_x B_{1-x} \rightarrow \frac{x - x_o}{1 - x_o} A + \frac{1 - x}{1 - x_o} A_{x_o} B_{1-x_o}, \tag{1.64}$$

leading to the line

$$\begin{aligned}\Delta H_f(A_x B_{1-x}) &= \left(\frac{1-x}{1-x_o}\right)\Delta H_f(A_{x_o} B_{1-x_o}) \\ &= \frac{\Delta H_f(A_{x_o} B_{1-x_o})}{x_o - 1} x - \frac{\Delta H_f(A_{x_o} B_{1-x_o})}{x_o - 1},\end{aligned}$$

which connects the stable phase at x_o with the the point at 0 concentration of B. If in Figure 1.2 there was a phase with a concentration of A between x and x_o, i.e. with two adjacent phases instead of a phase and a single elemental, then, if it is unstable, it would decompose in the two adjacent stable phases and the reasoning to determine its stability would be the same.

In general, the decomposition of a phase of concentration x in two other phases of concentrations x_2 and x_1 ($x_2 > x_1$) is given by

$$A_x B_{1-x} \rightarrow \frac{x - x_1}{x_2 - x_1} A_{x_2} B_{1-x_2} + \frac{x_2 - x}{x_2 - x_1} A_{x_1} B_{1-x_1} \tag{1.65}$$

and the line that determines the stability is

$$\Delta H_f(A_x B_{1-x}) = \frac{\Delta H_f(A_{x_2} B_{1-x_2}) - \Delta H_f(A_{x_1} B_{1-x_1})}{x_2 - x_1} x + \\ + \frac{\Delta H_f(A_{x_1} B_{1-x_1}) x_2 - \Delta H_f(A_{x_2} B_{1-x_2}) x_1}{x_2 - x_1}.$$

1.6 PHONONS

The structures resulting from MH runs are then "refined" by a geometry optimization, performed with stricter convergence parameters. This procedure sets the atomic positions and lattice parameters for which the forces over the ions and stresses over the cell are lower than some tolerance value. That is, the final structure is in a stationary point of the PES, i.e. in a point where the first derivative of the potential energy of the system (the forces on the atoms) vanishes. To distinguish whether such stationary points correspond to minima or saddle points, it is useful to look at the phonon modes, i.e. the frequencies of the modes of small vibrations of the atoms. If all the frequencies are real, the geometry corresponds to a minimum, and the structure is dynamically stable. Conversely, imaginary frequencies are indicative of structural instabilities.

Phonons are calculated by expanding in Taylor series the potential over the atoms around their equilibrium positions as a function of their displacements and retaining the second-order term. Consider a $N = N_x N_y N_z$ supercell of a periodic solid containing σ atoms and let $\xi^x_{\boldsymbol{R}\alpha}$ be the x coordinate of the displacement of the atom at position α in a unit cell located at position $\boldsymbol{R} = m_1\boldsymbol{R}_1 + m_2\boldsymbol{R}_2 + m_3\boldsymbol{R}_3$. The classical Hamiltonian describing the small vibrations of the atoms is

$$H_{\text{cl}} = \frac{1}{2} \sum_{\alpha \boldsymbol{R} x} \left[M_\alpha (\dot{\xi}^x_{\boldsymbol{R}\alpha})^2 + \sum_{\alpha' \boldsymbol{R}' x'} C^{xx'}_{\alpha\alpha'}(\boldsymbol{R} - \boldsymbol{R}') \right] \xi^x_{\boldsymbol{R}\alpha} \xi^{x'}_{\boldsymbol{R}'\alpha'}. \tag{1.66}$$

The term $C^{xx'}_{\alpha\alpha'}(\boldsymbol{R}-\boldsymbol{R}')$ is the mixed second-order derivative of the potential with respect to the displacement. It is called the interatomic force constant. The corresponding Hamilton equation is

$$M_\alpha (\ddot{\xi}^x_{\boldsymbol{R}\alpha})^2 + \sum_{\alpha' \boldsymbol{R}' x'} C^{xx'}_{\alpha\alpha'}(\boldsymbol{R} - \boldsymbol{R}') \xi^{x'}_{\boldsymbol{R}'\alpha'} = 0. \tag{1.67}$$

Because of the translation symmetry of the system, the solution can be looked for in the form

$$\boldsymbol{\xi}_{\boldsymbol{R}\alpha}(\boldsymbol{q}) = \boldsymbol{a}_\alpha(\boldsymbol{q}) e^{i[(\boldsymbol{q}\cdot\boldsymbol{R})+\omega_{\boldsymbol{q}}t]}. \tag{1.68}$$

To satisfy the periodic boundary conditions

$$\boldsymbol{q} = 2\pi \sum_{i=1}^{3} \frac{\nu_i}{N_i} \boldsymbol{G}_i, \tag{1.69}$$

where the $\boldsymbol{G}_i$ are the reciprocal lattice vectors and ν_i are integers satisfying the inequalities

$$-\frac{1}{2}N_i < \nu_i \leq \frac{1}{2}N_i \qquad (i = x, y, z). \tag{1.70}$$

The wavevector $\boldsymbol{q}$ can take N different values in the first Brillouin zone. As $N \to \infty$, $\boldsymbol{q}$ changes continuously. The vectors $\boldsymbol{a}_\alpha(\boldsymbol{q})$ characterize the direction of the vibrations in (1.68) with wavevector $\boldsymbol{q}$. Substituting (1.68) into (1.67), the Cartesian components of these vectors are determined as the solutions of the set of equations

$$\sum_{\alpha' x'} \widetilde{C}^{xx'}_{\alpha\alpha'}(\boldsymbol{q}) \boldsymbol{a}^{x'}_{\alpha'}(\boldsymbol{q}) - \omega^2_{\boldsymbol{q}} M_\alpha \boldsymbol{a}^x_\alpha = 0, \tag{1.71}$$

where

$$\widetilde{C}^{xx'}_{\alpha\alpha'}(\boldsymbol{q}) = \sum_{\boldsymbol{R}'} C^{xx'}_{\alpha\alpha'}(\boldsymbol{R} - \boldsymbol{R}') e^{i\boldsymbol{q}\cdot(\boldsymbol{R}-\boldsymbol{R}')} \tag{1.72}$$

form an Hermitian matrix. The eigenfrequencies ω are determined from the condition that the set of equations (1.71) has a solution

$$|\widetilde{C}^{xx'}_{\alpha\alpha'}(\boldsymbol{q}) \boldsymbol{a}^{x'}_{\alpha'}(\boldsymbol{q}) - \omega^2_{\boldsymbol{q}} M_\alpha \delta_{xx'} \delta_{\alpha\alpha'}| = 0. \tag{1.73}$$

As $\widetilde{C}^{xx'}_{\alpha\alpha'}$ is Hermitean and as the equilibrium position corresponds to an energy minimum, all 3σ roots of this equation will be real and positive functions of $\boldsymbol{q}$. Three among the 3σ frequencies vanish when $\boldsymbol{q}$=0 (which corresponds to an in-phase displacement of all atoms belonging to one elementary cell). These solutions, for any $\boldsymbol{q}$ are called acoustic branches modes. The remaining $3(\sigma - 1)$ frequencies are non-zero for q=0. These solutions are called optical modes.

The computation of the interatomic force constant can be generally carried out using a finite displacement technique or by a perturbative approach in the framework of Density Functional Perturbation Theory (DFPT), as described in References [39, 40]. In the first approach, the derivatives are determined by making small displacement of the atoms in the positive and negative directions along x, y and z, and computing the

corresponding energies. From Eequations (1.69) and (1.70), it can be seen that it is necessary to create a supercell to obtain a phonon band structure for non-zero vectors $\boldsymbol{q}$. In the DFPT approach, the derivatives are calculated using linear response in the reciprocal space by considering the change in the self-consistent potential in response to small atomic displacement.

Chapter 2

Binaries Phase Diagrams from First-Principles Structure Search

We present a methodology for the computation of binary phase diagrams from first-principles based on the Minima Hopping method and its application for the intermetallic binary alloys of sodium-gold and lithium-aluminum, and for the crystals of hydrogen-chlorine at moderate pressures. Our calculations unveiled several new stable phases in the three phase diagrams. A detailed electronic structure study of the alloys was conducted. The results for the Li-Al alloys were crosschecked with different echange-correlation energy functionals. We also examined the elastic properties by calculating the Bulk, Shear and Young moduli from *ab initio* elastic constants. We will then discuss the effect of pressure and temperature on binary phase diagrams, using the example of hydrogen-chlorine binaries. We found that triatomic molecules of hydrogen can be stabilized under moderate pressure in chlorine cages, while they were predicted to become stable in pure hydrogen above 2.1 TPa.

2.1 Introduction to Alloys

Alloys are materials composed of two or more elements, at such concentrations that an element can not be considered as an impurity. A typical purpose for making alloys is to enhance mechanical properties of the parent elementary solids as ductility, stiffness and hardness. For instance, the stiffness of aluminum can be enhanced by adding lithium, obtaining at the same time a decrease of wight. This fact makes aluminum-lithium alloys very attractive for the aerospace industry.

There are many examples of alloys in our everyday life. Among the most well known alloys are steel, which is composed of iron with a low content of carbon and stainless steel which also contains chromium. Iron is a very brittle metal, which is likely to brake if it receives sudden shocks. Adding carbon, one obtains a new alloy with higher toughness and tensile strength. Steel can be used in the construction industry for bridges and buildings. In the same way, adding chromium to iron increases the resistance to corrosion. Bronze is mainly composed of copper with the additions of

tin, although it might contain also nickel. Varying the amount of tin in the mixture, produces alloys that can have higher electrical conductivity, hardness, resonant qualities and resistance to the corrosion by sea water, as well as lower metal-metal friction than pure copper. Bronze has countless applications from musical instruments to automobile industry. Gold is mostly present in alloy form. The amount of gold in alloy is expressed in karats. One karat means that one out of twenty-four parts of the alloys is gold. Many metals can be used for the remaining portion of the gold alloys.

The more we understand alloys, the more efficiently they can be designed to target desirable features and properties. The earlier approaches to alloy composition estimation used phenomenological rules and experience in general. Today we can combine this empirical knowledge with first-principles computational methods. Computer solutions have progressively gained space and confidence in guiding material research, not only in academy but also in industry, thanks to the fact that computers have rapidly became more powerful and the methods more accurate. *Ab initio* methods offer a significantly access to the electronic and atomic structure of crystalline materials at a microscopic level that often is not directly accessible in experiment. Nowadays, in this domain, experiments and *ab-initio* calculation complement each other on an equal footing. Particularly, in the design of alloys, predictive computational methods are extremely useful, providing important information to experimentalists.

We can cite for example, the work of Ikehata and coworkers of the Toyota Central Research Development Laboratories, on the development of low elastic modulus titanium alloys [41]. These are light and high-strength materials that have been used in various engineering fields, from automobile and aerospace applications to biomaterials for load-bearing implants such as hip or knee prostheses because of their biocompatibility. Raabe and coworkers [42] used DFT calculations to screen between different possibilities of alloying titanium (Ta, Mo, Nb, and Zr) in order to obtain stable body centered cubic β-phase Ti-alloy satisfying the medical constraints of being nontoxic and having a reduced elastic stiffness. Sandlobes and coworkers [43] found that Yttrium, Scandium and several lanthanides in $Mg_{15}X$ solid-solution crystals, improved the room-temperature ductility of pure magnesium by up to 4–5 times which was confirmed in experiment. The industrial use of magnesium is hindered by its poor formability at room temperature, however, magnesium-based alloys are attractive for structural applications owing to their low mass density, good castability and efficient recyclability. Olson and coworkers [44] use DFT and other computational techniques to design ultra-high strength and high-toughness steels. Their DFT calculations indicate,

among other results, that in the bcc-iron crystal the gliding induced separation is dominant, which explains why plastic deformation takes place.

In sections 2.2 and 2.3, we will present the results of our calculations on alloys. Another example of alloy, in this case a semiconducting alloy, is discussed in section 3.3, for applications in photovoltaics.

2.2 Alloys of Sodium and Gold

Intermetallic compounds made of alkali metals and gold have intriguing electronic and structural properties that have not been extensively explored. We summarize here a systematic study of the phase diagram of one binary system belonging to this family, namely Na_xAu_{1-x}, using the ab initio Minima Hopping structure prediction method. We obtain that the most stable composition is $NaAu_2$, in agreement with available experimental data. We also confirm the crystal structures of $NaAu_2$ and Na_2Au, that were fully characterized in experiments, and identify a candidate ground-state structure for the experimental stoichiometry NaAu. Moreover, we obtain three other stoichiometries, namely Na_3Au_2, Na_3Au, and Na_5Au, that turn out to be thermodynamically stable in our simulations. Finally, we perform phonon calculations to check the dynamical stability of all reported phases and we simulate X-ray diffraction spectra for comparison with future experimental data.

Unusual and intriguing electronic properties have been found in binary phases of alkali metals and gold [45–48]. Proceeding through the alkali metal series from LiAu to CsAu, a metal-insulator transition occurs upon going from KAu to RbAu. All compounds have an ionic character, with gold being the negative ion. The structural properties of these binaries are also peculiar: while RbAu and CsAu adopt an octo-coordinated CsCl structure, analogous to the one of CsBr and CsI ionic crystals, the structures of the metallic NaAu and KAu compounds are more complicated and still not fully characterized. A new three dimensional icosahedral quasicrystal has recently been discovered by exploring the ternary Na-Au-Ga system [49].

Experimental studies on NaAu solid compounds date back to the beginning of the 20th century. Zintl *et al.* obtained a black deposit from an ammonia solution of sodium by adding Au, which they claimed to be a NaAu compound [50]. Since then, three other stoichiometric compounds were obtained from the melts of Na and Au, namely Na_2Au, $NaAu_2$, and $NaAu_5$ [46, 51]. While the crystal structure of Na_2Au and $NaAu_2$ were fully characterized, no definitive structure is given in literature for NaAu and $NaAu_5$ [46, 52–54]. As gold is a highly electronegative element, while alkali metals

are electropositive, one expects some ionic character for alkali metal-gold compounds. The negative oxidation state of gold is a focus of intensive studies as summarized in recent review papers [55, 56]. From the theoretical side, there are very few *ab initio* investigations of this family of compounds [57] and almost all the published calculations are devoted to NaAu clusters [47, 58, 59].

A further motivation to study NaAu binary systems comes from the fact that previous results on clusters and thin-films [45, 47, 58–60] hint at the existence of other stoichiometries, eventually stabilized by pressure [48].

Determining from first principles the phase diagram of a solid starting solely from its elemental composition is one of the most fundamental problems in materials science. The phase diagram of a binary compound represents the thermodynamic equilibrium of phases of a two-component system and it reveals useful insights into fundamental thermodynamic properties. However, the experimental determination of a phase diagram is an extremely time-consuming process, that requires careful synthesis and characterization of all possible phases. Accurate *ab initio* computational tools can significantly accelerate the construction of phase diagrams at zero temperature and zero pressure. Furthermore, it is straightforward in this calculations, to include the effects of pressure and temperature.

In this context, we use the Minima Hopping (MH) method [25, 26] coupled to Density Functional Theory (DFT) for the theoretical determination of the phase diagram of Na-Au. For a given stoichiometry, the initial geometries for the MH runs were obtained randomly, assuring only that the minimal distance between the atoms was at least equal to the sum of the covalent radii. We used cells containing up to 12 atoms, which, depending on the stoichiometry, corresponds to 1-4 formula units. During the structure search runs, forces and energies were obtained in the framework of DFT using the all-electron projector augmented wave method as implemented in VASP [37, 38]. The Perdew-Burke-Ernzerhof approximation (PBE) [12] to the exchange-correlation functional was used, and numerical convergence to less then 2 meV/atom was ensured by a high energy cutoff and dense k-point meshes.

We studied 19 different stoichiometries of $Na_xAu_{(1-x)}$, ranging from pure gold to pure sodium. As discussed above, excluding the pure elemental crystals, only four compositions were reported experimentally [61]: Na_2Au[46, 52], $NaAu_2$[46, 53], NaAu [46, 50], and $NaAu_5$ [46]. The structures of the first two are well known. A sample of Na_2Au of overall composition 69.95 at.% at 25°C Na was indexed as belonging to the $I4/mcm$ space group, crystallizing in a $CuAl_2$-type structure [46] with $a = 7.417$Å and $c = 5.522$ Å. On the other hand, $NaAu_2$ has $Fd\bar{3}m$ symmetry, and crystallizes in a

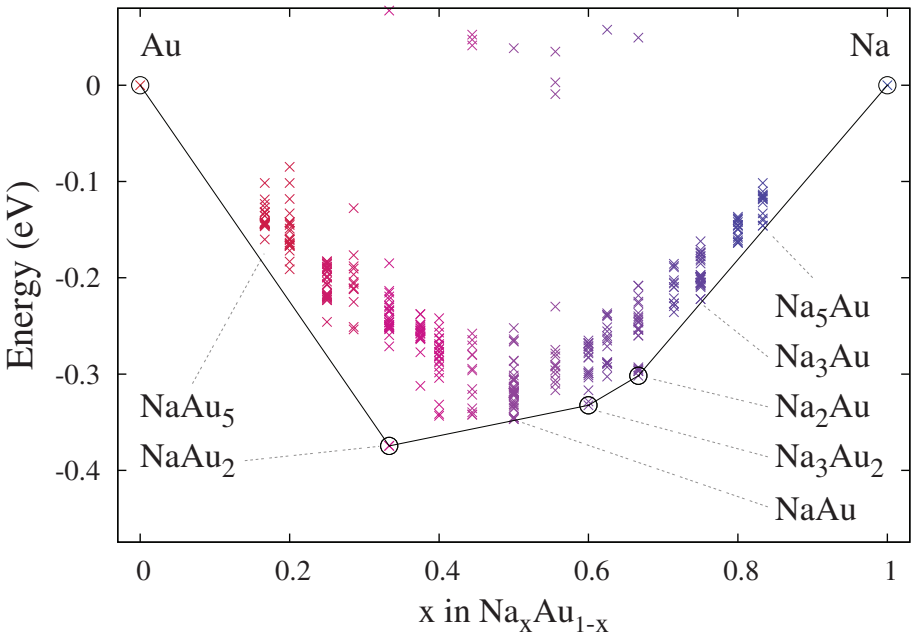
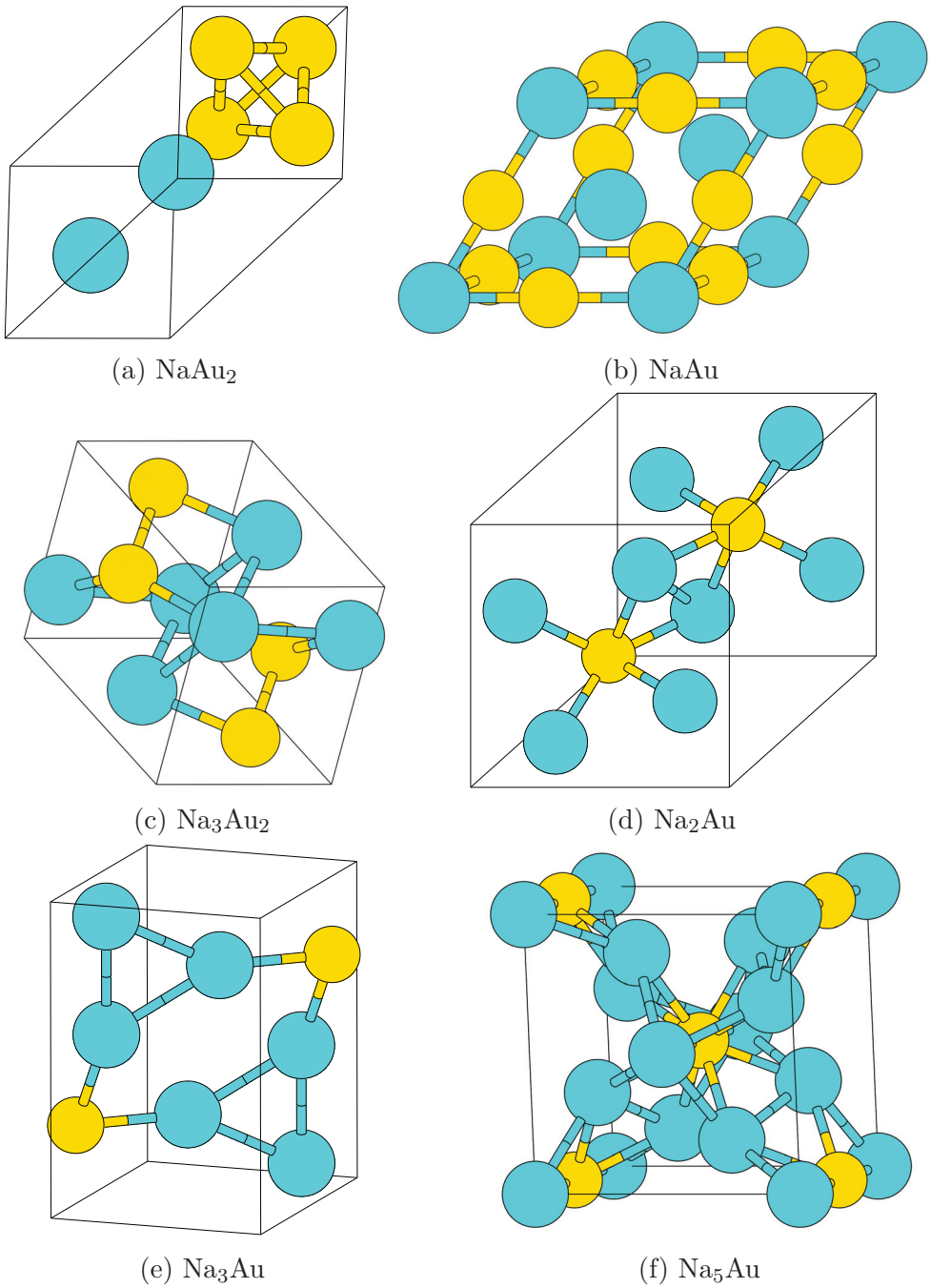


Fig. 2.1 Convex hull of the Na-Au binary system.

$MgCu_2$-type structure [46] with $a = 7.818$ Å (averaging on three compositions between 33 and 38 at.% at 25°C). NaAu has been reported to be a complex structure and no crystallographic data are given [46]. Finally, there exist some experimental evidence of the production of $NaAu_5$ at 775°C [46], which has not been further verified yet.

Figure 2.1 shows the convex hull of NaAu binaries, calculated considering all the lowest-energy structures resulting from the MH simulations for the different compositions $x = \mathrm{Na}/(\mathrm{Na} + \mathrm{Au})$. In this figure, low-energy structures are indicated by crosses, while thermodynamically stable compositions are marked with a circle. As discussed in the first chapter, the position of a structure with respect to the convex hull gives the primary criterion for its stability: all phases corresponding to points on top of the convex hull are stable, and those above are unstable. There are however, two details that we have to keep in mind: (i) There is always an error in the estimation of the formation energy inherent to our theoretical method. Therefore, structures that are "unstable" but close (∼10 meV/atom) to the convex hull can be stable experimentally. (ii) Structures that are close to the stability line can sometimes be stabilized by temperature, pressure, doping, or even intrinsic defects.

Based on these rules, we selected our lowest-energy phases which form the convex hull, or that are very close to it, for further theoretical characterization. The selected



(a) $NaAu_2$ (b) NaAu

(c) Na_3Au_2 (d) Na_2Au

(e) Na_3Au (f) Na_5Au

Fig. 2.2 Schematic representation of the structures of the Na-Au binaries found by the MH runs. Na is represented by blue spheres and Au by yellow spheres.

crystal structures are presented in Figure 2.2 and Table 2.1. Complete crystallographic information can be found in Appendix A.

It is worth to remark that the structure prediction runs yield the correct ground-state for both pure sodium ($P6_3/mmc$ structure with $a = 3.77\,\text{Å}$ and $c = 6.15\,\text{Å}$)

Table 2.1 Lowest-energy phases obtained from the MH calculations. Distances and angles in Å and degree respectively.

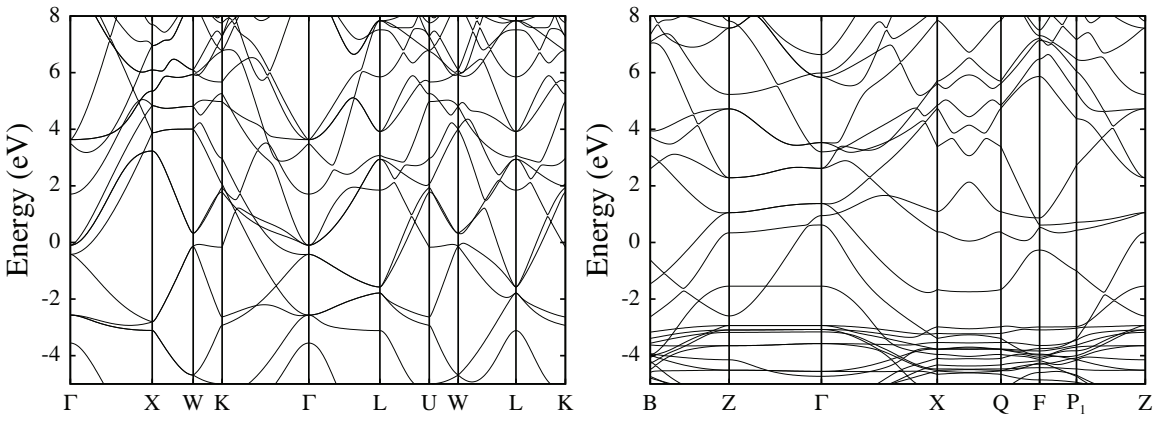
Phase	Space group	a	b	c	α	β	γ
$NaAu_2$	$F\,4_1/d\,\bar{3}\,2/m$	7.9473	7.9473	7.9473	90.00	90.00	90.00
NaAu	$R\,\bar{3}\,2/m$	5.6542	5.6542	14.3409	90.00	90.00	120.00
Na_3Au_2	$C\,1\,2/c\,1$	7.8593	8.5125	7.0438	90.00	95.15	90.00
Na_2Au	$I\,4/m\,2/c\,2/m$	7.3823	7.3823	5.6288	90.00	90.00	90.00
Na_3Au	$P\,1\,2_1/m\,1$	5.7324	9.2625	4.5781	90.00	113.29	90.00
Na_5Li	$P\,4/n\,2/b\,2/m$	7.6310	7.6310	6.2395	90.00	90.00	90.00

and gold ($Fm\bar{3}m$ structure with $a = 4.08$ Å). Furthermore, we also obtained the experimentally known structures of the Na_2Au and $NaAu_2$ phases. The theoretical lattice parameters of Na_2Au are $a = 7.38$ Å and $c = 5.63$ Å, and for $NaAu_2$ we find $a = 7.95$ Å. All lattice parameter are in very good agreement (of less than 2%) with the experimental values for Na_2Au, $a_{exp} = 7.417$ Å and $c_{exp} = 5.522$ Å and for $NaAu_2$, $a_{exp} = 7.818$ Å [46].

The proximity of the lowest energy phase for NaAu to the convex hull line suggests that this structure is a good candidate for the NaAu experimental sample. The NaAu structure has R$\bar{3}$m symmetry, with $a = 5.65$ Å and $c = 14.3$ Å. Of course, we cannot exclude that even more stable structures may be found using larger supercells in the structural search. Despite our extensive search, we could not find any stable structure for the other possible experimental stoichiometry, $NaAu_5$. Also in this case, more formula units per cell may be needed to identify a stable crystalline arrangement. However, our lowest-energy structure for $NaAu_5$ is quite far from the convex hull, raising the doubt that this stoichiometry might not be stable at all. Indeed, there is only one reported experimental observation, which has not been verified up to now. The analysis of the electronic band structures confirm that all the phases are metallic (see Figure 2.3).

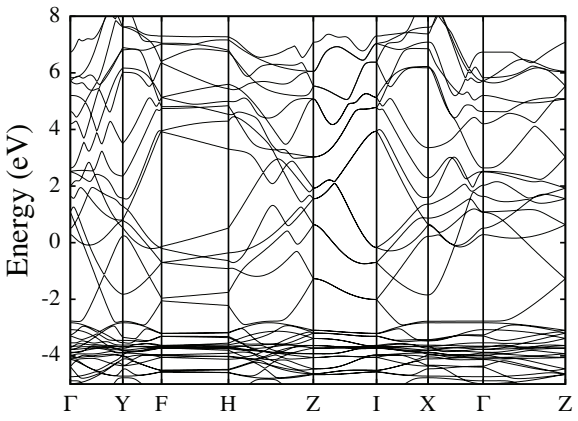
We also found three other structures on top or very close to the convex hull which were not reported experimentally: Na_3Au_2, Na_3Au, and Na_5Au. Na_3Au_2 crystallizes in a monoclinic $C2/c$ structure with $a = 7.86$ Å, $b = 8.51$ Å and $c = 7.04$ Å. Na_3Au has monoclinic $P2_1/m$ space group with $a = 5.73$ Å, $b = 9.26$ Å and $c = 4.58$ Å. Finally, Na_5Au has a tetragonal $P4/nbm$ structure with $a = 7.63$ Å, and $c = 6.24$ Å.

In order to further evaluate the stability of the new phases found by the MH calculations we evaluated the phonon spectra. In fact, the existence of imaginary frequencies in the spectra indicates dynamical instability, i.e., that the structure is

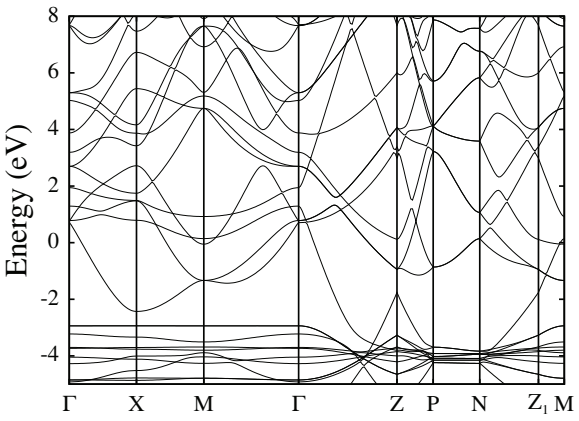


(a) $NaAu_2$

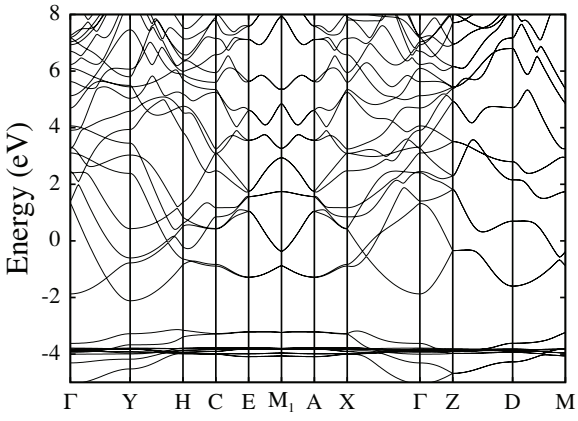
(b) NaAu



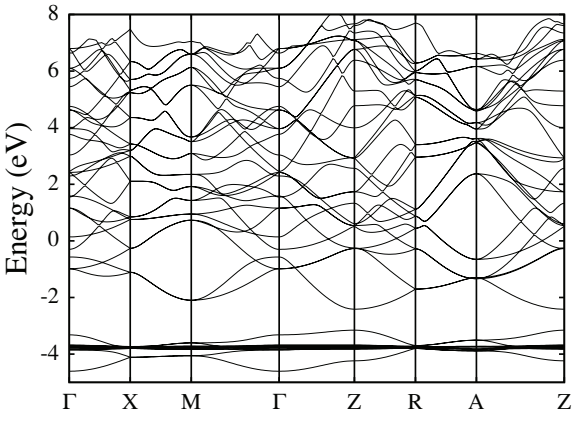
(c) Na_3Au_2



(d) Na_2Au



(e) Na_3Au



(f) Na_5Au

Fig. 2.3 Electronic band structures. The Fermi energy has been shifted to zero

Table 2.2 Bader charges on Na and Au inequivalent atoms for the lowest-energy structures. The Bader charges are given in elementary charge units

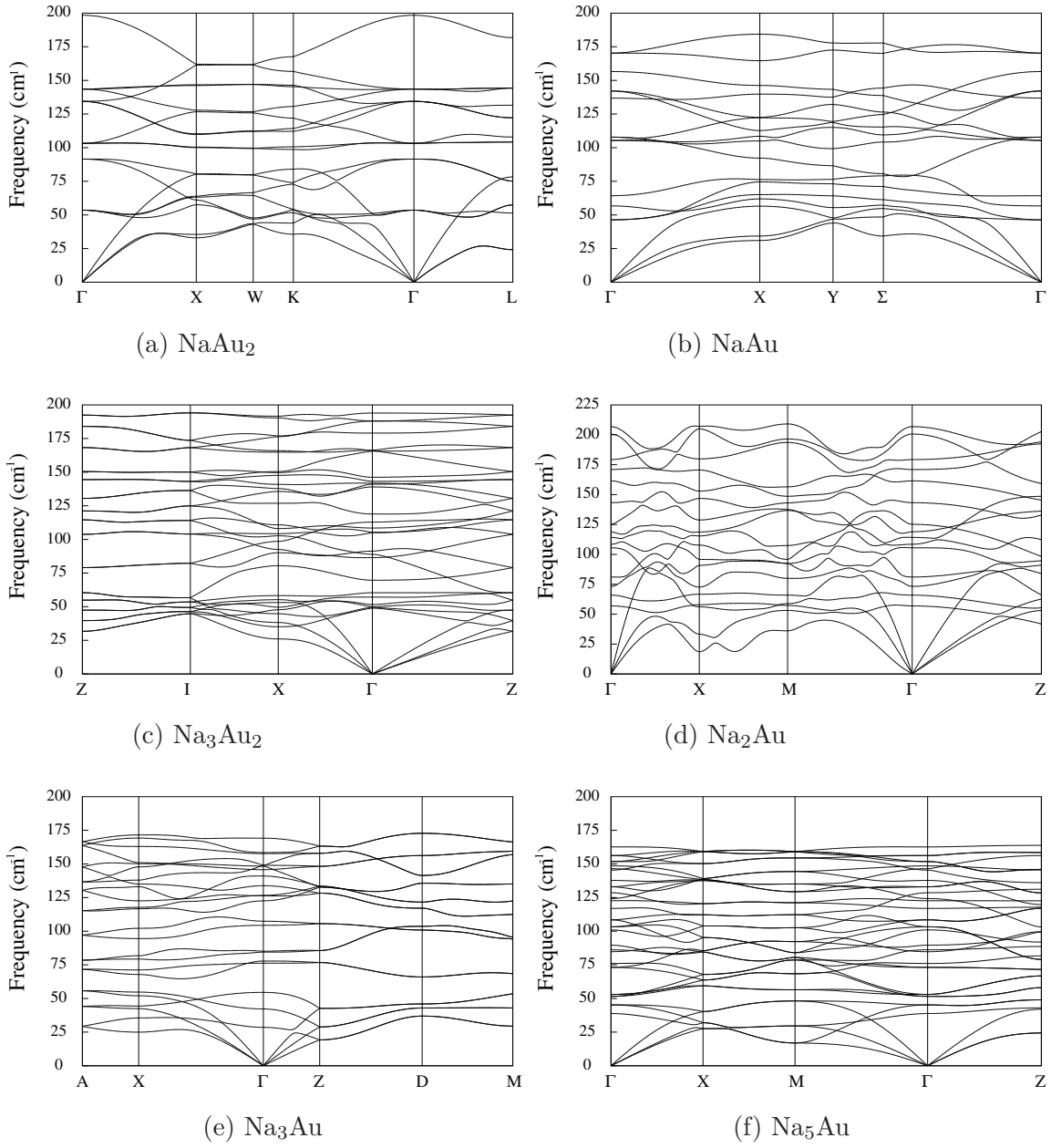
Stoichiometry	Charge on Na	Charge on Au
$NaAu_2$	+0.81	-0.40
NaAu	+0.70, +0.74	-0.72
Na_3Au_2	+0.71	-1.06
Na_2Au	+0.63	-1.25
Na_3Au	+0.60, +0.66	-1.85
Na_5Au	+0.18, +0.52	-2.25

not a minimum and that the correct structure contains a larger number of atoms in the unit-cell. The absence of imaginary frequencies is not a proof that the structure is the absolute minimum, but only indicates that it is a valid (meta-)stable phase. Our phonon calculations were performed with Density-Functional Perturbation Theory (DPT) [62–64] as implemented in the ABINIT [65] DFT software package. For the 6 structures shown in Figure 2.2 we used a cutoff energy of 40 Ha, a $2 \times 2 \times 2$ q-grid, and $4 \times 4 \times 4$ (for Na_5Au and Na_3Au_2) or a $8 \times 8 \times 8$ (for the other compounds) k-point grid.

Figure 2.4 shows the phonon band structures for the stable Na-Au phases we found. We can observe that, as expected, acoustical frequencies go to zero at Γ and there are no unstable modes present, clearly showing the dynamical stability of these phases. Furthermore, the maximum phonon frequency ranges from 160–190 cm^{-1}. These values are larger than the values for pure sodium ($\sim$120 cm^{-1} [66]) and pure gold ($\sim$150 cm^{-1} [66, 67]), indicating a stronger (ionic) bonding in the alloy.

The metallic character of these compounds and their larger phonon frequencies bring up the question of superconductivity in Na_xAu_{1-x}. According to the strong-coupling theory of electron-phonon mediated superconductivity, the transition temperature is proportional to the average phonon frequency. It is well known that neither sodium nor gold are superconducting at ambient pressure, but it is unclear if the increased phonon frequencies in Na–Au alloys is enough to yield superconductivity. To investigate this matter we performed linear-response calculations of the electron-phonon coupling with ABINIT. Our results indicate that these alloys are very likely superconductor, but with transition temperatures as low as 2 K due to the relatively small value of the electron-phonon coupling constant λ.

In Table 2.2 we report the Bader charges on the crystallographically non-equivalent atoms for the lowest-energy structures of each stoichiometry. We can clearly see that



(a) $NaAu_2$ (b) NaAu

(c) Na_3Au_2 (d) Na_2Au

(e) Na_3Au (f) Na_5Au

Fig. 2.4 Phonon band structures of the Na-Au phases.

the negative charge on gold increases monotonically by increasing the sodium content. In particular, it reaches the remarkable value of -2.25 for Na_5Au, showing that the Na-Au binary system is a perfect playground to study how gold behaves in very different chemical environments.

Finally, in Figure 2.5 we provide simulated X-ray diffraction spectra for the lowest-energy structures of the six stable compositions. Comparing with the experimental spectrum from Ref. [68] (reproduced in Figure 2.6), it is possible to notice that both

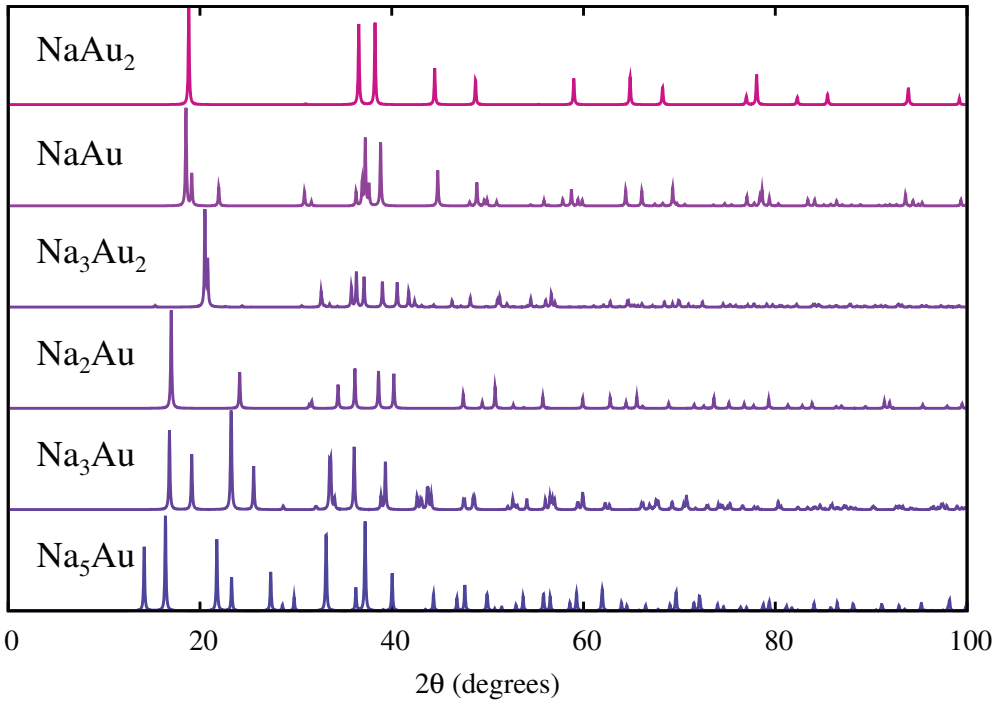


Fig. 2.5 Simulated X-ray diffraction spectrum with Cu Kα radiation $\lambda = 1.54178$Å.

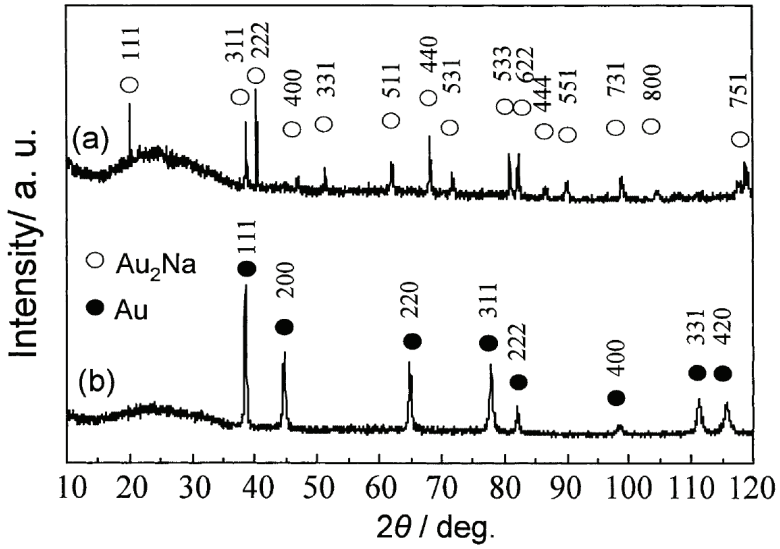


Fig. 2.6 X-ray diffraction patterns of Au electrode reproduced from Ref [68]: (a) after potentiostatic cathodic electrolysis at 0.3 V for 2 h; and (b) after potentiostatic anodic electrolysis at 0.6 V for 1 h after forming the Au_2Na alloy in a LiF-NaF-KF eutectic melt at 773 K. Au_2Na: cubic cell unit, a=0.779 nm.

$NaAu_2$ and NaAu agree well with experiments considering the large amount of noise present in the experimental data. In fact, the peak around $2\theta = 20°$ and $2\theta = 40°$ are present in the spectra of $NaAu_2$ and NaAu.

In summary, using MH we found two already known and four novel structures of Na_xAu_{1-x} that are likely to be thermodynamically stable. In the cases of Na_2Au and $NaAu_2$ our MH runs converged to the reported experimental geometries. We found a likely candidate for the crystal structure of NaAu, which was reported experimentally but not fully characterized. We also found for the first time a tetragonal Na_5Au phase, and monoclinic phases for Na_3Au_2 and Na_3Au. We calculated phonon band structures for the lowest-energy structures and we found that all of them are dynamically stable. Of course, the methodology presented here to study thermodynamic phase diagrams is absolutely general, and can be used for any other intermetallic binary as we shall see in the next section.

2.3 Alloys of Lithium and Aluminium

Intermetallic Li-Al compounds are, on one hand, key materials for light-weight engineering, and, on the other, they have been proposed for high-capacity electrodes for Li batteries. We determine from first-principles the phase diagram of Li-Al binary crystals using the MH structure prediction method. Beside reproducing the experimentally reported phases (LiAl, Li_3Al_2, Li_9Al_4, $LiAl_3$, and Li_2Al), a structural variety larger than expected is unveiled by discovering six unreported binary phases likely to be thermodynamically stable. Finally, we discuss the behavior of the elastic constants and of the electric potential profile of all Li-Al stable compounds as a function of their stoichiometry.

Light weight and good mechanical properties, together with attractive manufacturing costs, set Al alloys with Li at the forefront of aerospace materials research. The search for improved alloys is still open and attracts global interest for widespread industrial applications. It is well known that in the Al-rich range of stoichiometries (Li concentrations between 5% and 25%), Li–Al alloys offer higher stiffness and superior strength-to-weight ratios [69–71] than pure Al. In fact, each additional weight percentage of Li reduces the density by about 3% while increasing the Young modulus by about 6% [70]. This effect is explained by the hardening which occurs through the precipitation of nanoscale particles of the binary in the Al matrix [72]. The strengthening nano-precipitates, known as the δ' phase, are thought to be made of metastable $LiAl_3$ with a cubic ($Pm\bar{3}m$) structure, which remains crystallographically

coherent with the parent Al fcc matrix with small lattice mismatch. Although some of the strengthening qualities originate from the precipitates acting as pinning centers for defects, it is proved by *ab initio* calculations that the stoichiometric $LiAl_3$ crystal has a Young modulus significantly larger than Al [73, 74].

Li-Al alloys have been investigated extensively, both experimentally and theoretically, and diverse applications have been proposed for this system. In literature one can find information on five binary phases. While the Al-rich side of the phase diagram contains compounds with attractive elastic properties, the Li-rich side has been explored for high-capacity electrodes for Li batteries [75]. The experimental data show that starting from a face-centered cubic (fcc) metal (space group $Fm\bar{3}m$) of pure Al, addition of Li produces fcc Li–Al solid solutions until reaching the stoichiometry of the metastable simple cubic $LiAl_3$ and then two stable binaries, the cubic ($Fd\bar{3}m$) LiAl and, already in the Li rich range, the rhombohedral ($R\bar{3}m$) Li_3Al_2. The structure of Li_3Al_2 is indeed compatible with the rhombohedral ($R\bar{3}m$) zero temperature structure of Li. However, going from Li_3Al_2 to pure Li one passes through orthorhombic ($Cmcm$) Li_2Al and monoclinic ($C2/m$) Li_9Al_4 [76, 77]. The fact that at room temperature the crystal structure of pure Li becomes body-center cubic (bcc, $Im\bar{3}m$), and that at T=0 K the above mentioned rhombohedral, the bcc and an fcc phase are separated by less than 3 meV per atom according to first-principles calculations, is a clear indication of the complexity of bonding schemes that is brought by Li [74].

Theoretical investigations have mainly focused on the properties of the experimental phases, and little has been done in order to identify new ones. Podloucky *et al.* [73, 78] studied Al-rich alloys using first-principles statistical-mechanics and electronic-structure approaches within the local density approximation [3] (LDA) to the exchange-correlation energy. They determined thermodynamical magnitudes considering a statistical ensemble of possible fcc supercells containing up to 8 atoms. Alam *et al.* [79] also tried fcc and bcc-based compounds without finding any new stable structure. Also, using the Korringa-Kohn-Rostoker method within the coherent-potential approximation (KKR-CPA), they predicted a higher stability of off-stoichiometry Al-rich phases with respect to the ordered ones, when the Li content is between 8% and 20%.

Despite this large amount of studies in literature, the knowledge on the whole phase diagram of Li–Al appears to be still incomplete. In view of that, this section presents a first-principles structure prediction study of ordered phases of the Al-Li binary system. We used the same methodology described in the previous section to investigate the phase diagram of Li_xAl_{1-x}.

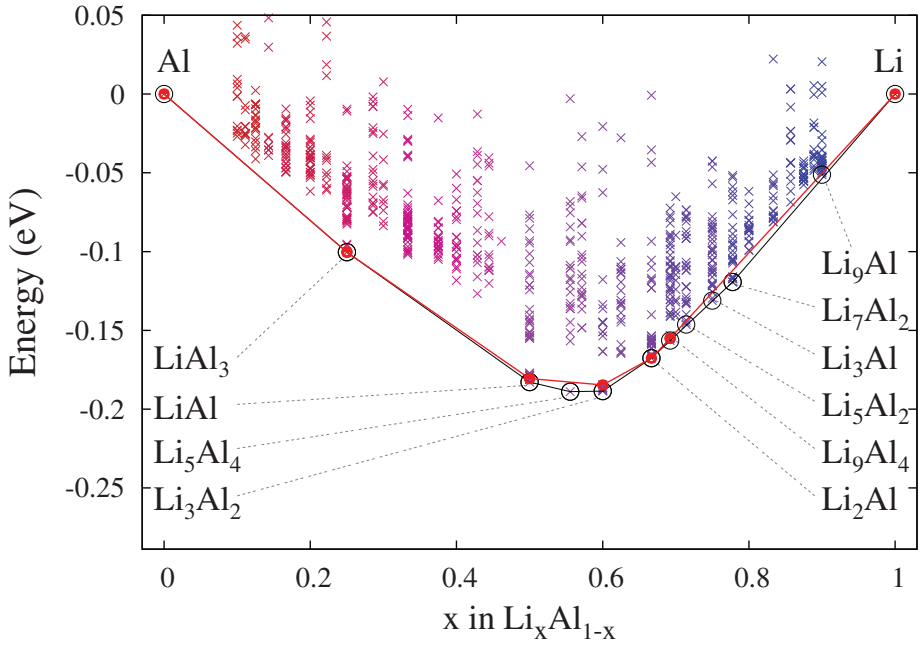


Fig. 2.7 Convex hull of the Li-Al binary system: one constructed from the *ab initio* MH calculations (black with open circles) and the other (red with filled circles) with structures from databases.

We explored 33 different stoichiometries from pure Al to pure Li. For a given stoichiometry, the initial geometries for the Minima Hopping [25, 26] (MH) runs were obtained randomly, ensuring only that the minimal distance between the atoms was at least equal to the sum of the covalent radii. We used cells containing up to 13 atoms, which, depending on the stoichiometry, corresponds to 1–5 formula units. Forces and energies were obtained in the framework of Density Functional Theory (DFT) using the all-electron projector augmented wave (PAW) method as implemented in VASP [37, 38]. We used the Perdew–Burke–Ernzerhof (PBE) approximation [12] to the exchange-correlation functional, together with a plane-wave cutoff of 520 eV and a number of k-points sufficient to guarantee a numerical convergence of total energies to less than 2 meV/atom.

With the lowest energies structures obtained from the MH calculations, we constructed the theoretical convex hull, which is presented in Figure 2.7 with a black line. To showcase the performance of our method in predicting crystal structures, we include in Figure 2.7 also the convex hull constructed with the energy of the previous known Li–Al phases (red line). Most of these were available through the AFLOWLIB [16] and OQMD [17] databases, while the structure of Li_2Al was taken from Refs. [76, 77]. The

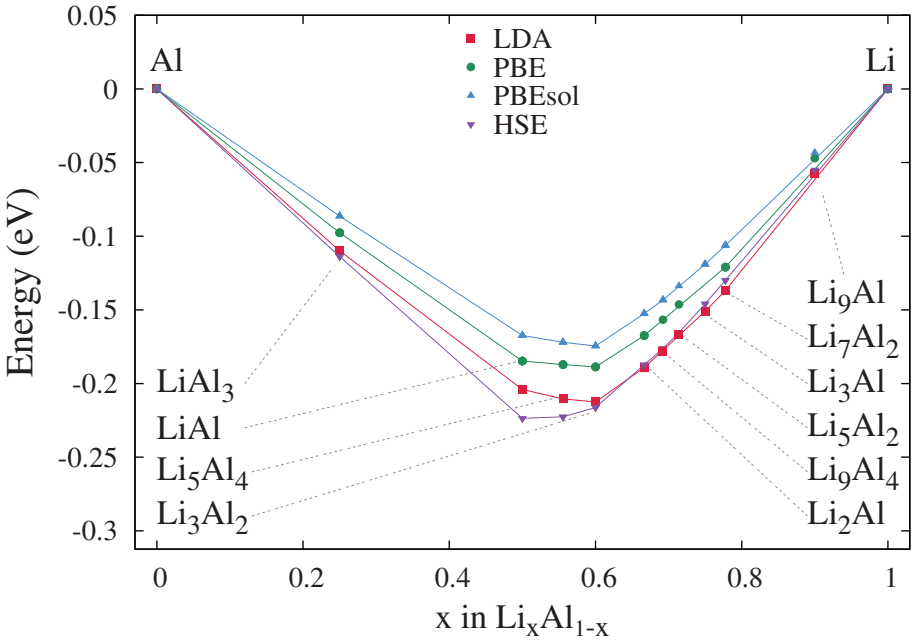


Fig. 2.8 Convex hull of the Li–Al binary system using different exchange-correlation functionals. Red line with plus symbols: computed with the LDA. Green line with crosses: computed with PBE. Blue line with squares: computed with PBEsol. Dark grey line with triangles: computed with HSE at the PBEsol geometries. The thermodynamically stable compositions are indicated by the symbols on the convex hull lines.

energy of the thermodynamically stable compositions, which form the convex hull, are marked with a circle. Red filled circles refer to experimental structures and the crosses indicate the energies of the low-lying local minima explored by the MH algorithm during our simulations. To select the stable compositions, we consider the structures with energies close (within 10 meV/atom) to the convex hull as belonging to it. In this way, we take into account effects like the inherent errors in the estimation of the formation energy, temperature, pressure and defects, which may slightly modify the position of the structures with respect to the stability lines.

In order to evaluate the effect of the choice for the exchange-correlation functional on the convex hull we repeated the reoptimization step using different approximations, i.e. the LDA [3], PBEsol [80] (a revised PBE functional that improves equilibrium properties of densely packed solids and their surfaces), and the screened hybrid Heyd-Scuseria-Ernzerhof (HSE) [81, 82] at the PBEsol geometries. The results are shown in Figure 2.8, were we can see that the shape of the convex hull is to a large extent

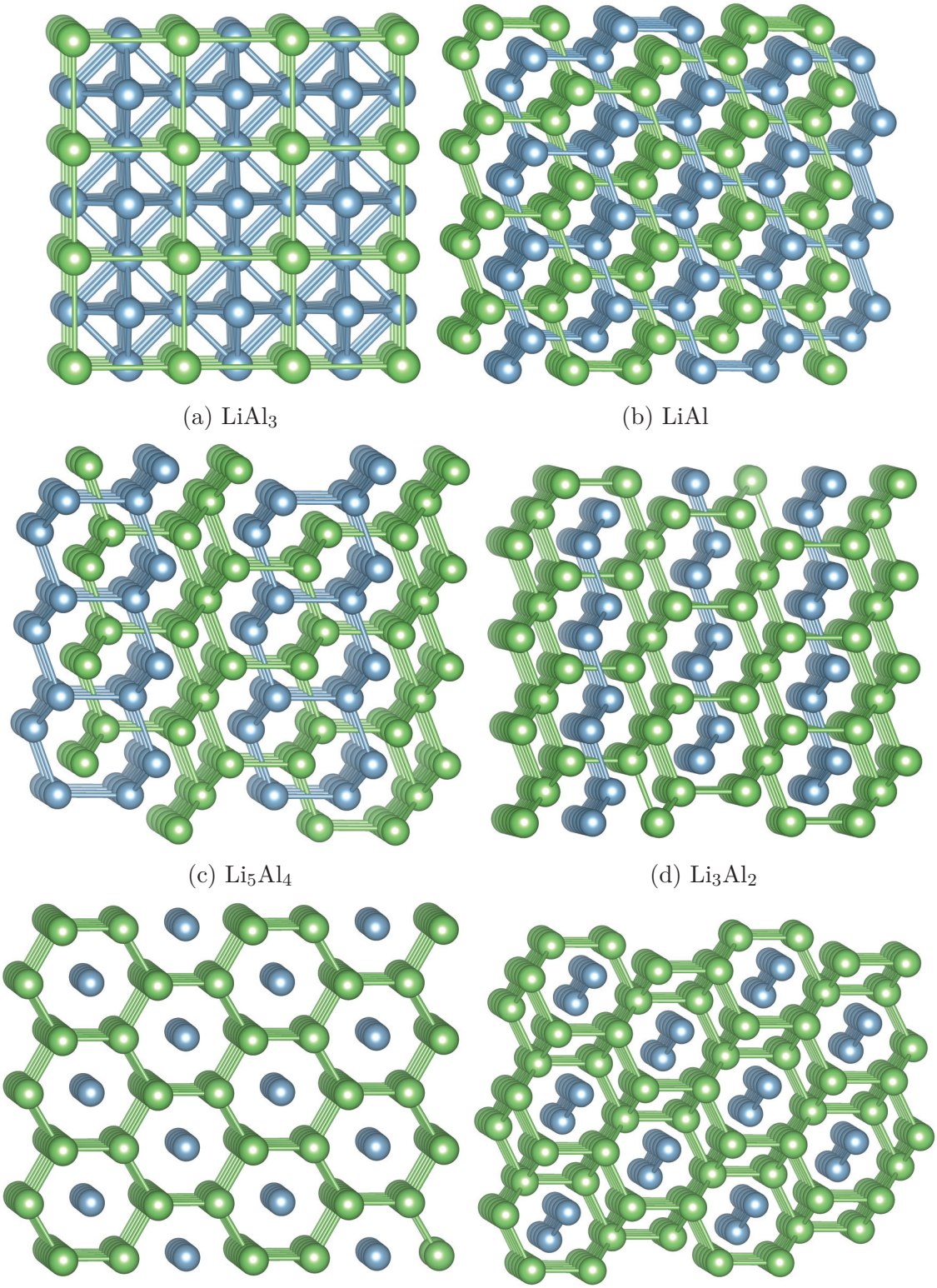
Table 2.3 Lowest-energy phases obtained from the MH calculations. Distances and angles in Å and degree respectively. (∗) Reported experimental phase of Li_2Al. Complete crystallographic information can be found in Appendix B.

Phase	Space group	a	b	c	α	β	γ
Al	$Fm\bar{3}m$	4.041	4.041	4.041	90.00	90.00	90.00
$LiAl_3$	$Pm\bar{3}m$	4.025	4.025	4.025	90.00	90.00	90.00
LiAl	$Fd\bar{3}m$	6.348	6.348	6.348	90.00	90.00	90.00
Li_5Al_4	$P\bar{3}m1$	4.446	4.446	8.391	90.00	90.00	120.00
Li_3Al_2	$R\bar{3}m$	4.446	4.446	14.073	90.00	90.00	120.00
Li_2Al	$Cmcm(*)$	4.592	9.596	4.448	90.00	90.00	90.00
Li_2Al	$P6/mmm$	4.541	4.541	2.667	90.00	90.00	120.00
Li_9Al_4	$C2/m$	5.367	4.453	18.633	105.41	103.14	90.17
Li_5Al_2	$C2/m$	5.401	4.465	9.698	90.00	95.27	90.00
Li_3Al	$P12_1/m1$	5.420	4.460	5.636	90.00	100.03	90.00
Li_7Al_2	$R\bar{3}m$	7.932	7.932	8.226	90.00	90.00	120.00
Li_9Al	$P\bar{1}$	5.076	5.093	7.936	83.45	71.71	70.30
Li	$R\bar{3}m$	3.067	3.067	22.085	90.00	90.00	120.00

independent of the choice for the exchange-correlation potential, and therefore the same compositions are thermodynamically stable for all the tested approximations.

The stable crystal structures are summarized in Table 2.3 and their atomic arrangements are displayed in Figures 2.9 and 2.10. We found eleven stable intermetallic structures, out of which six had not been reported before. The other five are the experimentally known structures mentioned above. For these, the lattice constants were determined in very good agreement of less than 2% with the experimental values [74, 76, 77, 83, 84]. It is important to stress that the MH algorithm was able to find all the experimentally known structures, and to determine new phases lower in energy than the ones available in the databases. This confirms its predictive power, which is especially relevant in this system that presents a large diversity of atomic arrangements.

From a careful analysis of Figure 2.7 we can notice that almost straight lines connect LiAl with pure Al, passing through $LiAl_3$, that touches the convex hull. This is in agreement with the known metastability of $LiAl_3$: since it lies on a straight line, it will likely coexist with pure Al and LiAl. The same occurs on the straight line connecting Li_3Al_2 and Li_7Al_2: the phases lying on the convex hull between these two compounds can easily coexist in mixtures of the two extreme compositions. We can conclude that LiAl, Li_3Al_2 and Li_5Al_4 are the most stables binary phases. Note that Li_5Al_4 is

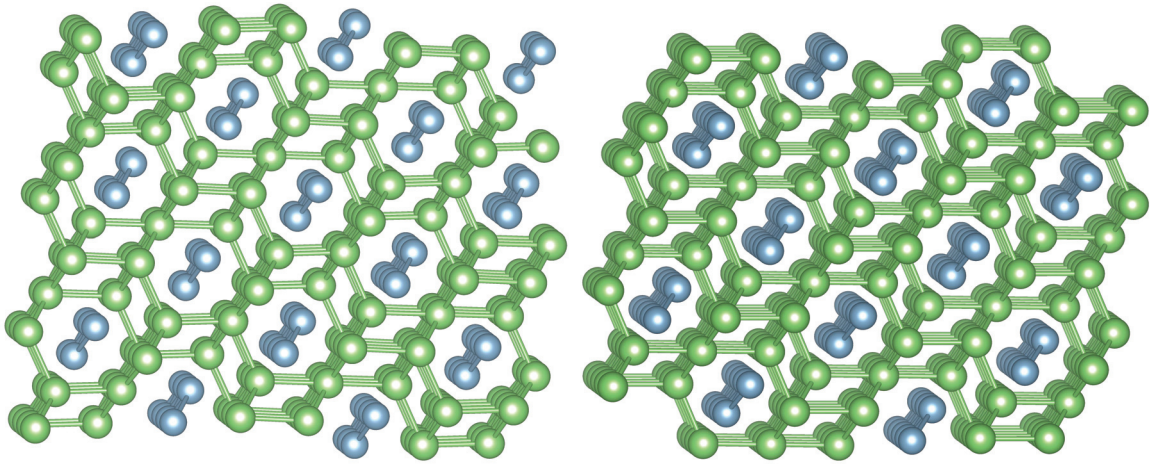


(a) $LiAl_3$ (b) LiAl

(c) Li_5Al_4 (d) Li_3Al_2

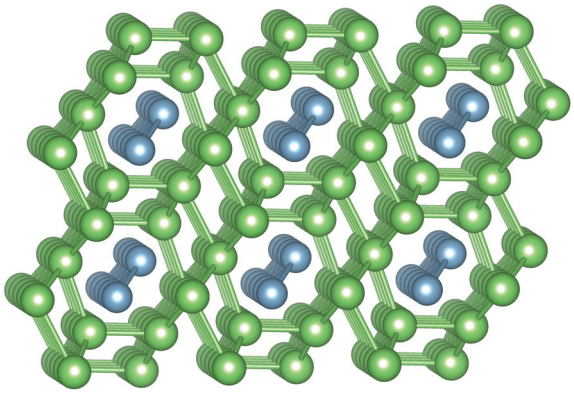
(e) Li_2Al (hexagonal) (f) Li_2Al (orthorhombic)

Fig. 2.9 Crystal structure of $LiAl_3$, LiAl, Li_5Al_4, Li_3Al_2, Li_2Al and Li_2Al. Aluminium, blue and lithium, green

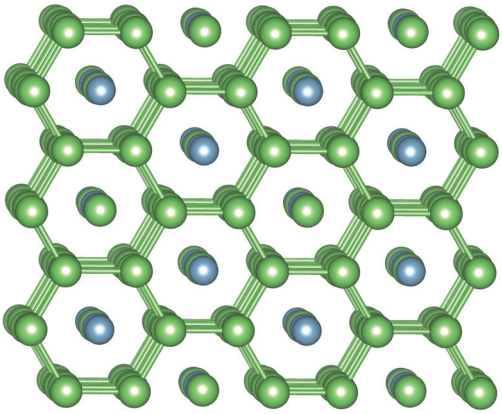


(a) Li_9Al_4

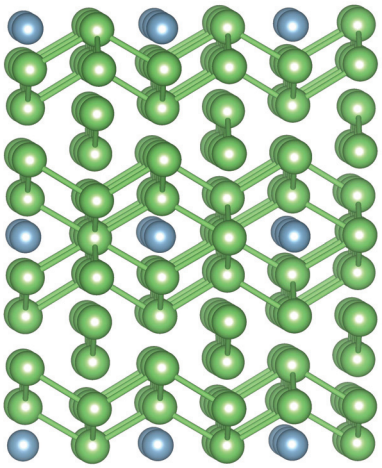
(b) Li_5Al_2



(c) Li_3Al



(d) Li_7Al_2



(e) Li_9Al

Fig. 2.10 Crystal structure of Li_9Al_4, Li_5Al_2, Li_3Al, Li_7Al_2 and Li_9Al. Aluminium, blue and lithium, green.

absent from known databases. Even if most of the Li–Al compounds that we identify as thermodynamically stable can easily decompose, they will likely exist as microscopic inclusions in a matrix of another stoichiometry, contributing therefore to the physical properties of the alloy. It is therefore relevant to study the structural, elastic, and electronic properties of the Li–Al lowest-energy binaries all over the composition range.

The two stable crystal structures of $LiAl_3$ and LiAl can be easily related to the original fcc Bravais lattice of Al. $LiAl_3$ is a simple cubic crystal with 4 atoms in the unit cell, where Li occupy the vertices of the cubes and Al the center of the cube faces, recreating the original fcc lattice but with two different atoms on the lattice sites. The Al-Al and Al-Li distances are 2.85 Å, basically unmodified with respect to fcc Al. We can therefore see that this crystal structure is perfectly compatible with the starting fcc lattice of Al. LiAl is made of two interpenetrating fcc lattices, one made of Li and the other one of Al, displaced by $(0.5\ 0.5\ 0.5)a$, where a is the length of the conventional cube. In this structure each atom has an equal number of Al and Li neighbors. Li-Li distances and Al-Al distances measure 2.75 Å, which is significantly smaller than the bond length in pure Al (2.86 Å) and pure Li (3.05 Å).

The Li rich side of Li–Al alloys is more interesting concerning the complexity of the ground-state structures, with intermetallic compounds that differ significantly from their parent pure crystals. On this side of the phase diagram we found several stable or almost stable compounds for a Li content greater than 50 % that were not reported in literature. We remind that Li has a rhombohedral $R\bar{3}m$ ground state structure at zero temperature characterized by a Li-Li bond length of 3.05 Å, but fcc and bcc structures are only few meV per atom higher in energy. This gives a hint that the structural variety can be much larger for Li-rich compounds. Li_9Al has a distorted bcc structure, with one out of 10 cube centers occupied by Li instead of Al. The bond length between Li atoms is basically unchanged with respect to pure Li. Li_7Al_2 and Li_3Al_2 are rhombohedral, Li_5Al_4 is trigonal, and Li_3Al has several orthorhombic and monoclinic phases within few meV from the hull.

Indeed, while we can observe on the Al rich side of the phase diagram that the lowest energy minima are clearly separated from the other geometrically distinct minima by sizable energy gaps, Li rich compounds show a quasi-continuous distribution of low-lying minima (see Figure 2.7). This suggests that Al rich ordered crystals with small unit cells should be easier to obtain experimentally than Li rich crystals. In the latter case one can rather expect at finite temperature the coexistence of different low-energy structures, possibly in a disordered phase.

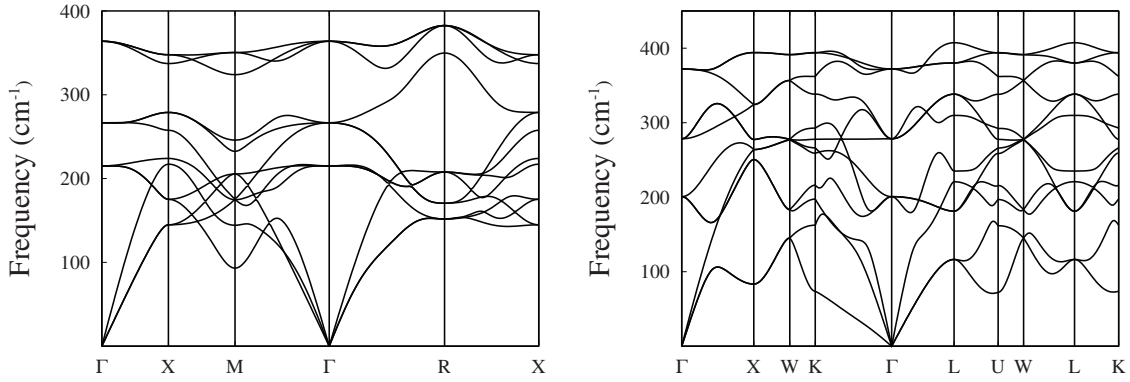
Concerning the stoichiometry Li_2Al, which is experimentally reported as an orthorhombic crystal [77] with space group $Cmcm$, we could find an hexagonal $P6/mmm$ and a trigonal $P\bar{3}m1$ structure which are only 2 and 5 meV/atom higher in energy, respectively. In the hexagonal phase Li atoms are arranged in aligned honeycomb lattice layers, forming hexagonal cells with one Al in the center. In the orthorhombic structure Al atoms are instead arranged in zigzag chains, while the trigonal phase contains zigzag planes of Al. Also in this case, all three structures (and relative mixed configurations) must be considered as possible stable phases due to the very small energy differences.

The structures with about 50 % of Li are peculiar as the bond lengths are very different with respect to both parent Li and Al compounds: Li-Li, Al-Al and Li-Al bonds are contracted compared to the neighboring compositions. In general, the Li-Li bond is reduced in all compounds on the Li-rich side except for Li_9Al. We can also observe that Al atoms tends to have other Al as nearest neighbors even for Li-rich stoichiometries [74].

To further study the stability of the new phases we also calculated phonon bands. Frozen phonons calculations were performed with the software PHONOPY from force constants calculated using VASP for 2x2x2 supercells with finite displacements. All forces were obtained within the LDA. Figures 2.11 and 2.12 show phonon band structures for the stable phases. No unstable modes were identified. Although this does not guarantee that the identified structures constitute the ground state at each composition, it validates them as dynamically stable. Maximum phonon frequencies range from 382 cm^{-1} to 447 cm^{-1}.

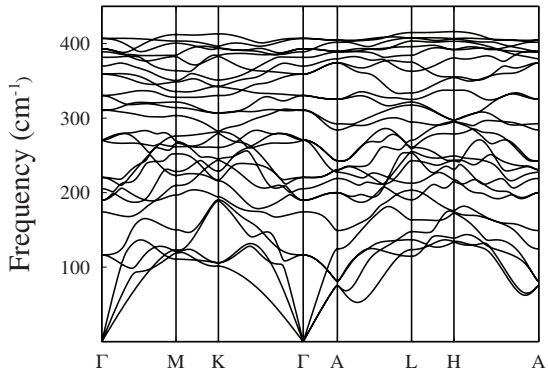
In view of the importance of the mechanical properties of Li–Al alloys, we was calculated the dependence of the atomic density and of the elastic constants on the Li content. As it is desirable to simulate polycrystalline samples, that can be macroscopically described as isotropic elastic materials, we used appropriate averages of the elastic constants. The behavior of polycrystalline linear elastic materials can be described with a set of two independent parameters. Here we select the bulk modulus B and the shear modulus G and we estimate them using the Voigt-Reuss-Hill (VRH) averaging scheme [85], which gives, respectively, upper and lower boundaries for the elastic constants [85]. If the values obtained in the two approaches are close, their arithmetic media is a reliable estimate of the searched quantity. In the Voigt scheme, the bulk and shear moduli can be estimated from the the elastic stiffness coefficients as

$$9B = (c_{11} + c_{22} + c_{33}) + 2(c_{12} + c_{23} + c_{31}) \tag{2.1}$$

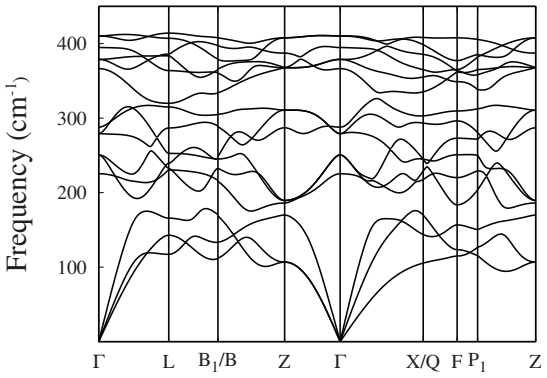


(a) $LiAl_3$

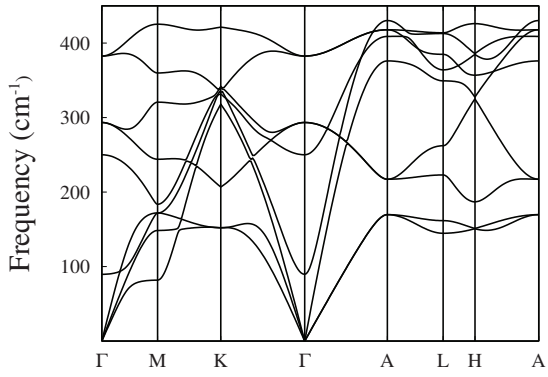
(b) LiAl



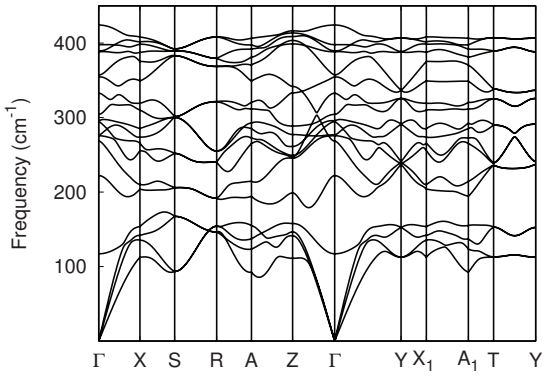
(c) Li_5Al_4



(d) Li_3Al_2

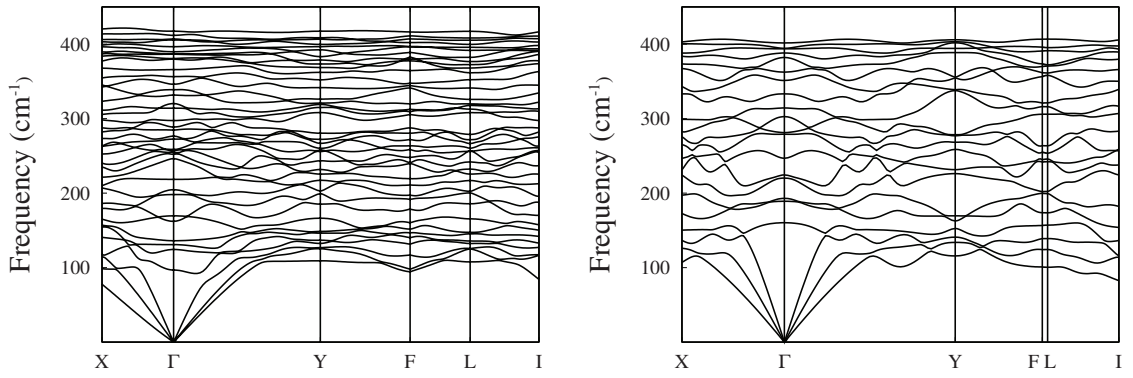


(e) Li_2Al (hexagonal)



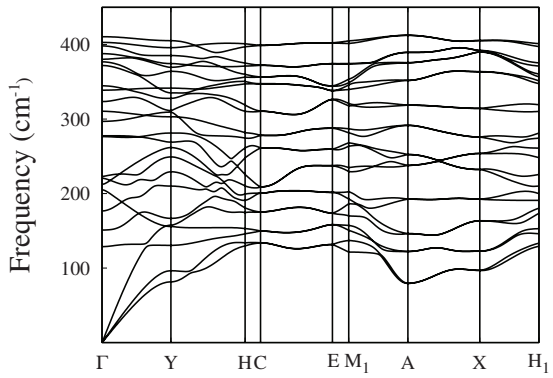
(f) Li_2Al (orthorhombic)

Fig. 2.11 Phonon band structures of $LiAl_3$, LiAl, Li_5Al_4, Li_3Al_2, Li_2Al and Li_2Al.

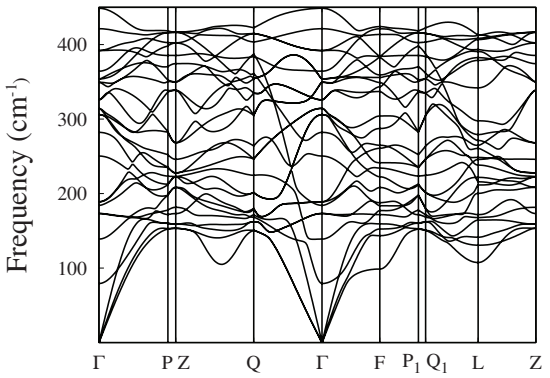


(a) Li_9Al_4

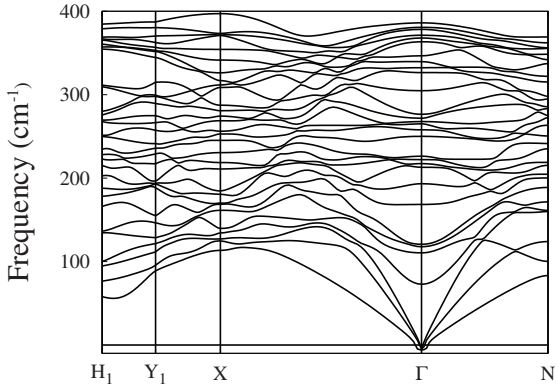
(b) Li_5Al_2



(c) Li_3Al



(d) Li_7Al_2



(e) Li_9Al

Fig. 2.12 Phonon band structures of Li_9Al_4, Li_5Al_2, Li_3Al, Li_7Al_2 and Li_9Al

Table 2.4 Calculated bulk (B), shear (G), Young (E) moduli in GPa, Poisson's ratio (ν) and B/G for the different phases. Experimental values are given for pure Li and Al.

Phase	B	G	E	ν	B/G
Al	78.1	30.5	80.9	0.33	2.6
exp. [88]	75.86	28.34	75.60	0.33	2.7
$LiAl_3$	63.4	39.8	98.7	0.24	1.6
LiAl	47.6	30.9	75.7	0.24	1.5
Li_5Al_4	41.8	33.6	79.0	0.19	1.2
Li_3Al_2	38.9	32.8	76.4	0.17	1.2
Li_2Al	34.9	33.1	75.4	0.14	1.1
Li_2Al	32.6	27.1	63.3	0.18	1.2
Li_9Al_4	31.8	25.4	59.8	0.19	1.3
Li_5Al_2	29.0	22.6	53.4	0.19	1.3
Li_3Al	27.8	21.6	51.2	0.19	1.3
Li_7Al_2	27.6	21.7	51.5	0.19	1.3
Li_9Al	18.4	9.4	23.5	0.29	2.0
Li (bcc)	12.9	5.5	14.0	0.32	2.3
exp. [89]	12.13	8.78	21.22	0.21	1.4

and

$$15G = (c_{11} + c_{22} + c_{33}) - (c_{12} + c_{23} + c_{31}) + 2(c_{44} + c_{55} + c_{66}), \tag{2.2}$$

respectively, while in the Reuss scheme, they are obtained from the elastic compliance coefficients

$$1/B = (s_{11} + s_{22} + s_{33}) + 2(s_{12} + s_{23} + s_{31}), \tag{2.3}$$

$$15/G = 4(s_{11} + s_{22} + s_{33}) - 4(s_{12} + s_{23} + s_{31}) + 3(s_{44} + s_{55} + s_{66}). \tag{2.4}$$

Only 9 out of the 21 independent elastic components of the elastic tensors for the single crystal are present in these expressions. It already has been noticed that, shear moduli determined from VRH averages are not always in good agreement with experimental values [84, 86]. Uesugui *et al.* [84] determined shear modulus using GGA, larger by 86.7% than experimental values. A similar situation is found for the Young modulus of LiIn, determined by Kuriyama *et al.* [87]. The Young modulus measured on a polycrystalline aggregate of InLi was larger by a 75% than the one determined from the VRH averages of elastic constants. Averaged bulk moduli are instead usually in good agreement with experiments.

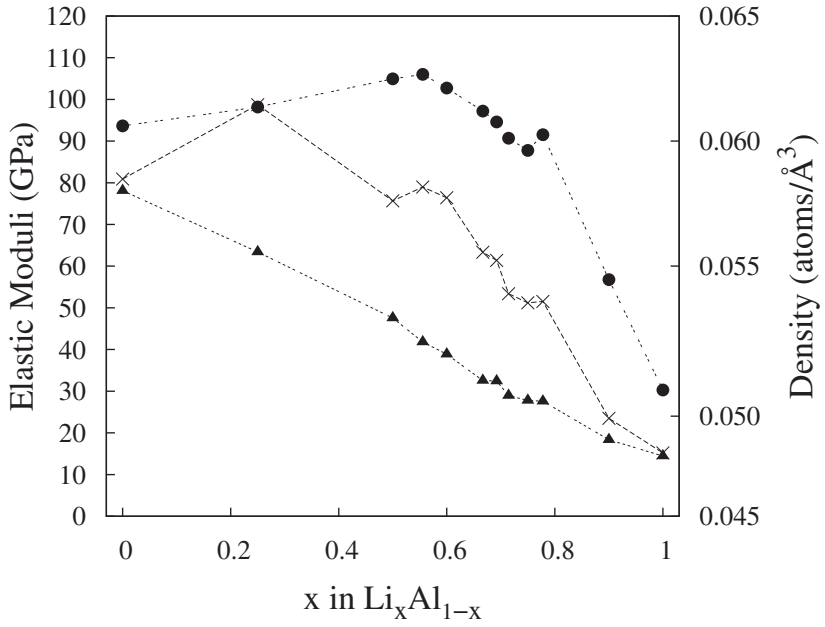


Fig. 2.13 Bulk modulus (filled triangles, in GPa), Young modulus (crosses, in GPa), and atomic number density (filled circles, in atoms/Å^3) as a function of the Li concentration

Once that B and G are determined, it is easy to calculate the Young modulus E and Poisson's ratio ν, using $G = E/(2+2\nu)$ and $B = E/(3-6\nu)$. The elastic constants were calculated from the stress and strain relations by performing six distortions of the lattice, as implemented in VASP [37, 38]. In order to assure the convergence of all quantities by less than 1 GPa, we had to increase the kinetic energy cutoffs to 1000 eV and the k-point meshes to $15 \times 15 \times 15$. The results are summarized in Table 2.4 and Figure 2.13, where we report the calculated values for the bulk modulus B, the shear modulus G, the Young modulus E and the Poisson's ratio ν of all stable phases. We remark that these calculations are performed using a PBE exchange-correlation functional. We also calculated the elastic moduli with the LDA. The variations of the values of the elastic constants are smaller than 10% (see Figure 2.14).

In Figure 2.13 we show the dependence of the density and the elastic moduli on the Li concentration. The calculated moduli are in reasonably good agreement with available experiments [88, 89]. The new structures meet the properties that were already reported: when increasing the amount of Li the bulk modulus [90, 91] decreases almost monotonically with the volumetric mass density. This is consistent with a much stronger resistance of Al to volume changes by applied pressure. However, the Young modulus increases on the Al-rich side of the phase diagram by small addition of Li,[69–71] in apparent contrast with the fact that the Young modulus of Li is 5 to 6

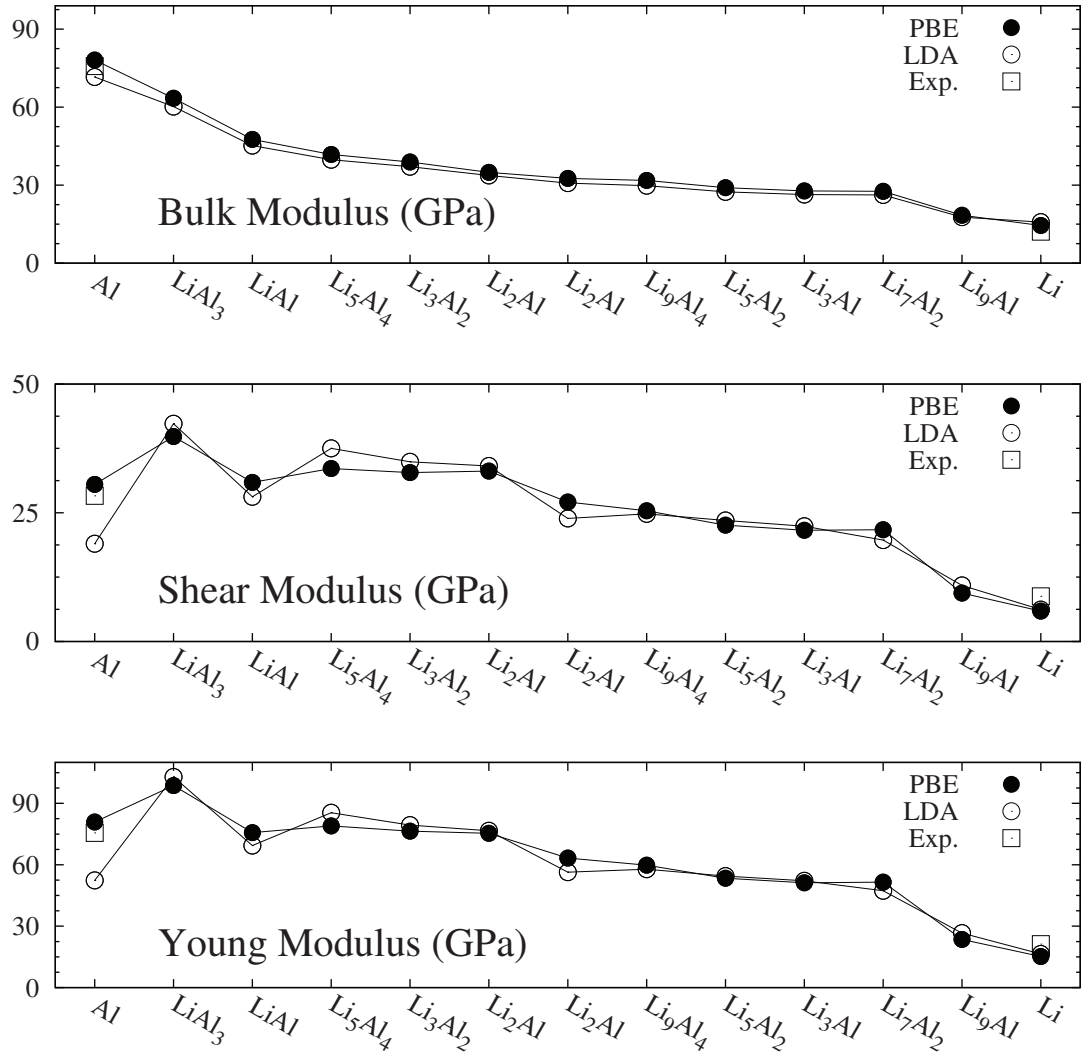


Fig. 2.14 Comparison of the calculated PBE and LDA bulk (B), shear (G), Young (E) moduli in GPa, Poisson's ratio (ν) and B/G for the different phases. Experimental values are given for pure Li and Al (see Table 2.4)

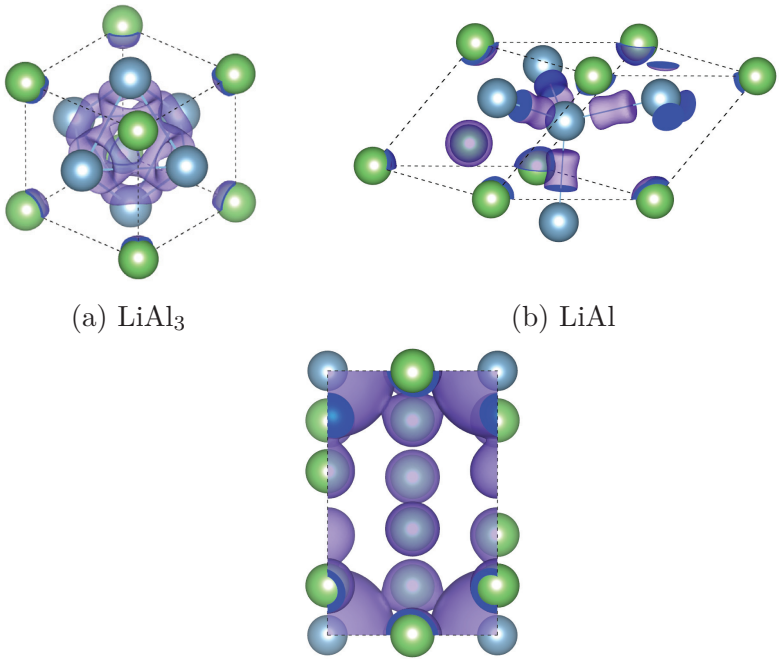
times smaller than the one of Al. We further notice that the Young modulus of $LiAl_3$ is considerably higher than that of pure Al, and that the Young modulus is essentially constant up to concentrations of 60% of Li, when it starts decreasing at a fast rate. A similar dependence on the composition is shown by the atomic density, which is also plotted for comparison in Figure 2.13. The Young modulus is together with the shear modulus a good estimate of the stiffness of the material. This shows that, from a purely mechanical point of view, alloys containing substantially more than 5-25% of Li could be used without deterioration of the stiffness.

Table 2.5 Bader charges on Li and Al atoms for the lowest-energy structures. The Bader charges are given in elementary charge units. (∗) The electron is in general distributed between neighbors Al atoms as shown in Figure 2.15. (∗∗) seven Li atoms with -0.2 and two Li atoms with +0.4.

Phase	Li	Al(∗)
Al	0.0	-
$LiAl_3$	-0.8	+0.3
LiAl	-0.8	+0.8
Li_5Al_4	-0.8	+0.9, +1.2
Li_3Al_2	-0.8	+1.2
Li_2Al	-0.8	+1.6
Li_9Al_4	-0.8	+1.8
Li_5Al_2	-0.8	+2.0
Li_3Al	-0.8	+2.4
Li_7Al_2	-0.8	+2.8
Li_9Al	-0.8,+1.4 (∗∗)	+3.8
Li	-	0.0

In the context of Li batteries the mechanical properties of the electrodes are also important to prevent mechanical failures as a consequence of the stress developed due to large volume expansion upon lithiation. The ratio between the bulk modulus and shear modulus gives an estimate of the ductile or brittle behavior of a material: B/G ratios smaller than 1.75 indicate brittle alloys, while B/G is large than 1.75 in ductile alloys. Most of the considered structures are brittle. Beside the obvious exception of pure Al, which is ductile, also Li_9Al and Li have a B/G ration larger than 1.75. Nevertheless, by comparison with the experimental value of B/G for Li (1.4), we understand that this is an artifact due to the coincidental overestimation of B and underestimation of G. These results point to the risk of mechanical failures in LiAl alloys when used for electrodes [92]. The Poisson's ratio ν is defined as the ratio of transverse strain to the longitudinal strain and it gives information on the resistance of the material against shear, therefore giving indirect access to bonding forces. The value of the ν is usually in the interval 0 – 0.5, and larger values of ν reflect better plasticity. Our results show an average plasticity with a better behavior while approaching pure Al and pure Li.

The phase stability and the bonding characteristics of the alloys can be better understood by looking at the distribution of the valence charge around each atom. A characteristic of Al-rich compounds is that Li atoms redistribute some of their valence



(a) $LiAl_3$ (b) LiAl

(c) Li_9Al

Fig. 2.15 Electronic density in in the unit cell of $LiAl_3$, LiAl, and Li_9Al. (aluminium, blue and lithium, green)

electrons to the Al bonds (see Figure 2.15) and the resultant strengthened Al bonds stabilize the compounds [74, 84]. We observed this redistribution of charge (except for Li_9Al where the quantity of Al is too small) by calculating the Bader charges for all the studied compounds (Table 2.5). This added stabilization can be easily seen in the plot of the atomic number density of the different compositions, measured in atoms per unit volume (Figure 2.13). As we see in Figure 2.13 there is an excellent correlation between the Young modulus and the atomic number density. (Note that this is in contrast with the usual volumetric mass density that decreases almost linearly as a function as increasing Li concentration). In fact, up to concentrations of 60% of Li, the increased bonding leads to denser and therefore stiffer structures. This effect is maximized for $LiAl_3$.

In summary, we have discovered several unknown ordered phases of the Al-Li binary system that are thermodynamically stable. Phonon band structures indicate that these new Li-Al phases are also dynamically stable. Analysis of the elastic constants indicate that the stiffness of LiAl alloys with up to 60% of Li remains essentially equal to the one of Al, with a marked maximum at $LiAl_3$. This can be understood by the stabilization

of these compounds (due to an electron transfer from Li atoms to Al bonds) that increases the atomic number density and therefore the stiffness. These results expand our knowledge of the Li–Al phase diagram and can have profound influence in the understanding and design of new Li–Al alloys for light-weight engineering and batteries.

In the next section we calculate phase diagrams including the effects of pressure and temperature for a very different system: H-Cl binaries.

2.4 Triatomic Hydrogen Molecules in H-Cl Crystals under Moderate Pressures

Triatomic hydrogen is a controversial molecule, whose existence was under debate during most of the XX^{th} century. In 1979, conclusive experiments proved that H_3 can be produced in a low pressure gas discharge tube [93]. However, it exists only in an excited state, with its third electron being delocalized far from the three protons, namely in a Rydberg state [94]. If the molecule attempts to lose energy and go to the repulsive ground state, it spontaneously breaks up. The lowest energy metastable state, has an energy lower by few eV than the H^+ and e^- state but decays in around 1 ps [95]. In this work we propose that H-Cl binaries under pressure are a viable way to access and characterize the elusive triatomic molecule of hydrogen in an experimental setup. The synthesis at moderate pressures and room temperature of these non-conventional stoichiometries of HCl could therefore considerably expand our knowledge of chemical bonds in hydrogen.

We used the Minima Hopping (MH) method to calculate the phase diagram of H-Cl binaries when an external pressure is applied. We allowed any stoichiometry H_xCl_y that was compatible with a combination of up to 10 atoms in the unit cell. The energies were calculated with VASP [37, 38] using the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [12]. At zero pressure, orthorhombic HCl is the only stable point of the H-Cl phase diagram. At moderate pressures other stoichiometries richer in hydrogen become stable. The convex hulls of H-Cl binaries at 50 GPa and 100 GPa are presented in Figure 2.16. We can see that one new composition (H_3Cl) is stabilized at 50 GPa and three additional compositions (H_3Cl, H_5Cl and HCl_2) are stable at 100 GPa.

We are mostly interested in the H-rich side of the phase diagram. H_3Cl, the first composition to become stable under pressure, touches the convex hull as low as about 10 GPa followed by H_5Cl, slightly above 80 GPa. The lowest-enthalpy structure of H_3Cl up to about 90 GPa belongs to the $C2/c$ space group and presents planar zig-zag

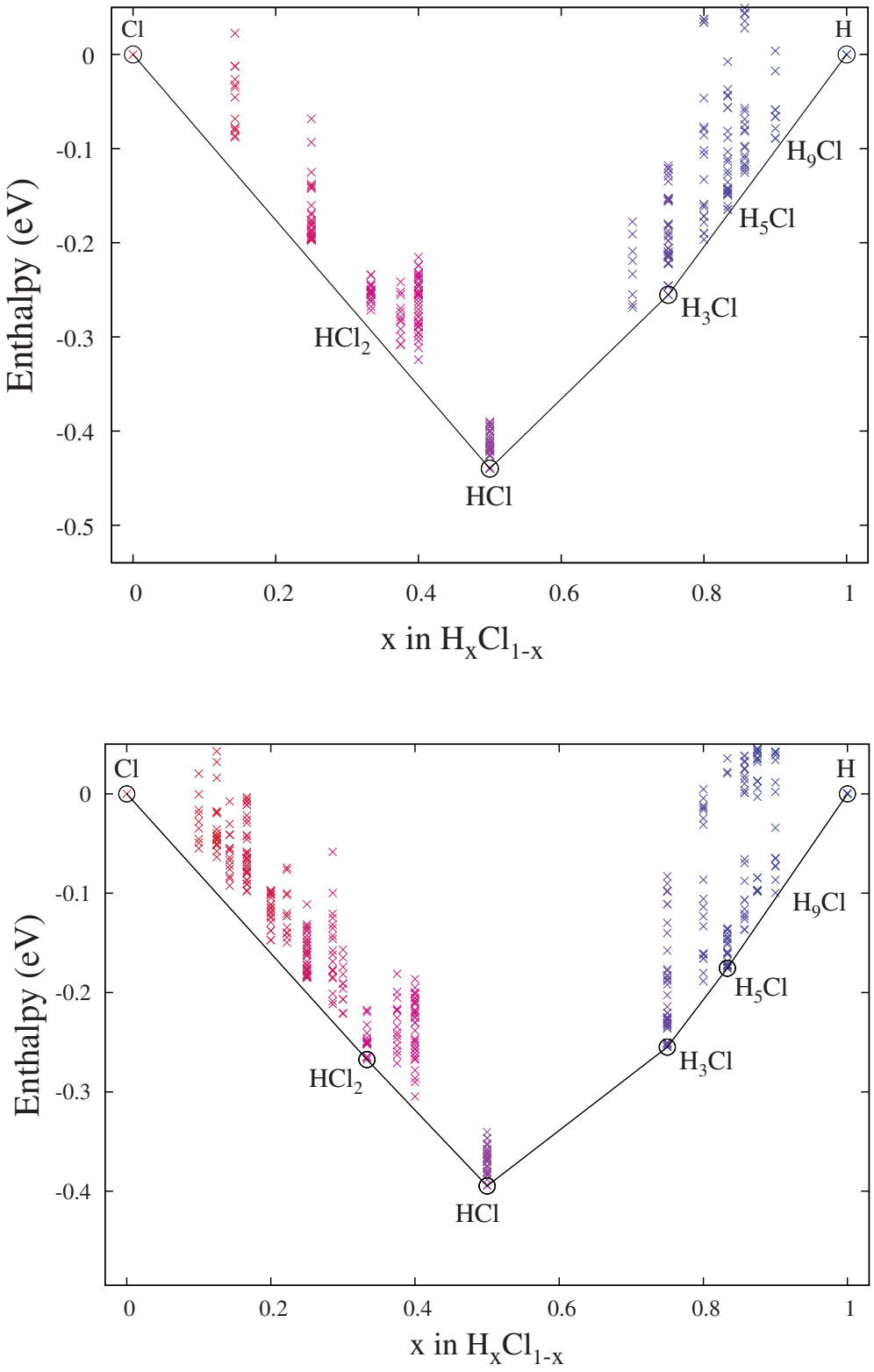


Fig. 2.16 Phase diagrams of the H-Cl system at 50 GPa (top) and 100 GPa (bottom).

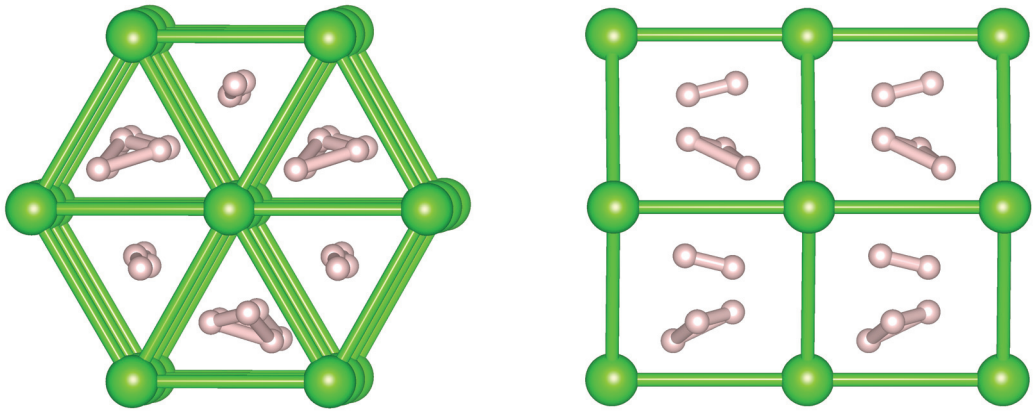


Fig. 2.17 Schematic representation of the structures of H_5Cl at 100 GPa along the c (left) and b (right) lattice vectors.

chains of HCl intercalated with H_2 molecules. At 86 GPa a phase with P-1 symmetry becomes the most stable: this crystal is very similar to the $C2/c$ phase, however it displays a different stacking and orientation of the chains/molecules, respectively. Finally, at around 180 GPa, a phase with a completely different structural motif emerges, belonging to the space group $Ama2$. The Cl atoms form a host structure with hexagonal channels surrounding H_3 triangular guest clusters. At a pressure of 150 GPa, the triangles are scalene with bond angles of 67.7°, 47.5°, 64.8°, which become more and more equilateral with increasing pressure (62.5°, 53.4°, 64.1° at 200 GPa).

A similar behavior is also observed for the stoichiometry H_5Cl. Below 82 GPa the lowest enthalpy phase of H_5Cl has P-1 symmetry and displays the previously discussed planar zig-zag chains. Two H_2 molecules per HCl unit are located in between the chains. At around 82 GPa the H-Cl chains disappear and a phase with Cc symmetry becomes the most stable, where the Cl atoms form hexagonal channels (see Figure 2.17). Inside these channels we find H_2 molecule, with the typical bond length of 0.74 Å, together with H_3 triangular clusters with sides of 0.8 Å - 1.1 Å. In analogy to H_3Cl, the triangles are scalene and become essentially equilateral with increasing pressure.

We also observed triangular clusters of hydrogen, contained together with H_2 molecules in Cl tetragonal cages, in the lowest enthalpy structures of H_9Cl. This phase is not thermodynamically stable yet at 100 GPa, even if it gets very close to the convex hull.

Figures 2.18 and 2.19 show the evolution of the phonon density of states of H_5Cl with pressure. At 0 GPa there are imaginary phonon modes which disappear when pressure is applied. The binary phase is unstable, it is a mixture of H_2 and HCl

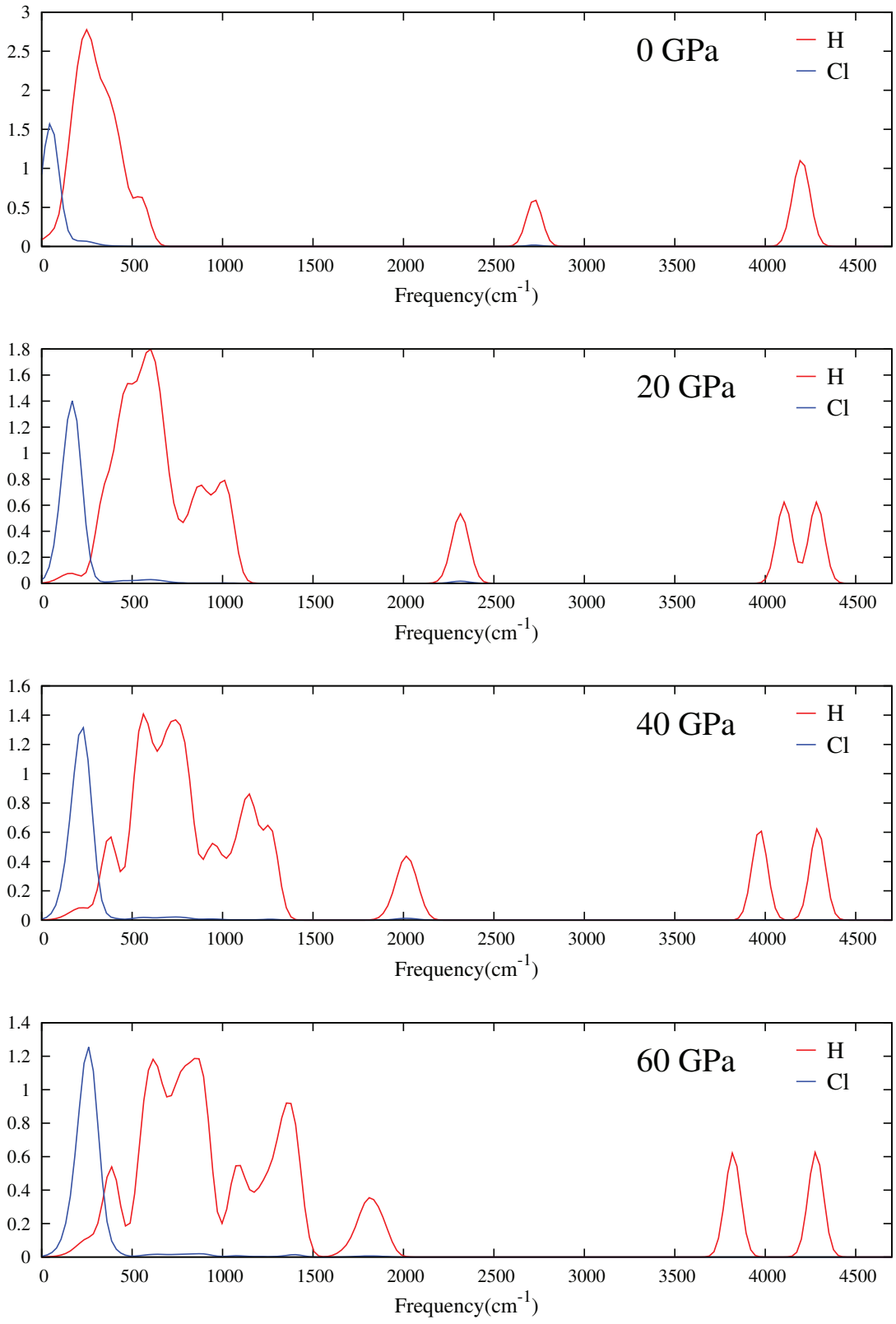


Fig. 2.18 Phonon density of states for the Cc phase of H_5Cl.

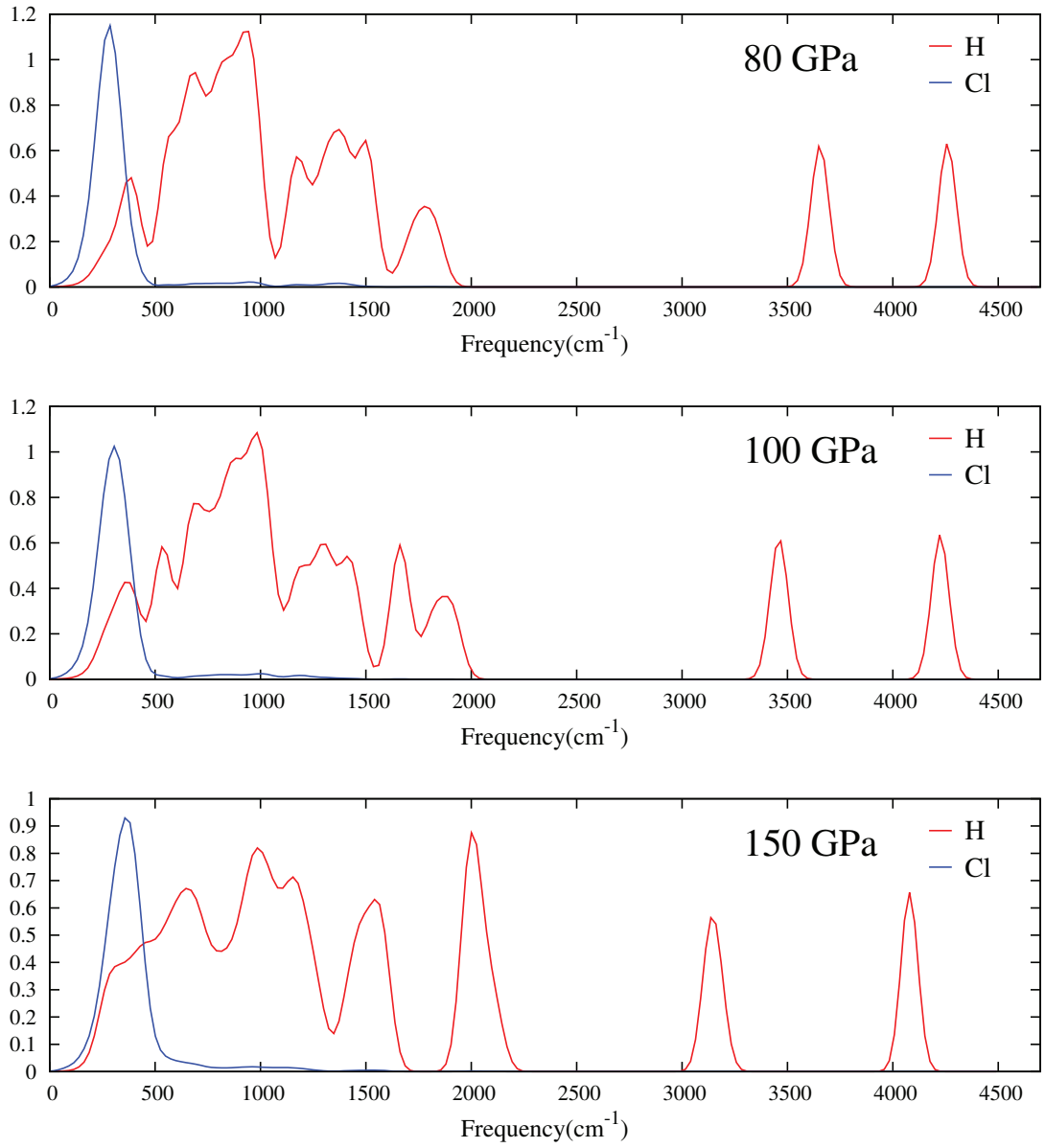


Fig. 2.19 Phonon density of states for the Cc phase of H_5Cl (continuation).

molecules. We can observe the contribution to the phonon spectrum of the normal modes of vibration of both molecules: HCl, with a bond length of 1.2987 Å, at about 2700 cm^{-1} (exp. 2991 cm^{-1}, exp. bond length 1.2746 Å [96]) and H_2, with a bond length of 0.75 Å [96] at about 4200 cm^{-1} (exp. 4401 cm^{-1}). As the pressure increases, it is visible a phase transition. Besides the stabilization of the imaginary modes that were present before, the modes corresponding to the vibration of the H_2 molecule splits in two and the modes corresponding to the vibration of the HCl molecule move to

a lower frequencies. This happens as, when the cell is compressed, the H_2 and HCl molecules get closer in a T-shape relative orientation (see Figure 2.21a) and the H-Cl bond stretches.

At 150 GPa, we can distinguish two peaks, one at ∼3100 cm^{-1}, which can be attributed to the breathing mode of the H_3 molecule (exp 3178.18 cm^{-1} [96]) and one at ∼2000 cm^{-1}, which could correspond to the tentatively assigned value of 2109.7 cm^{-1} [96, 97] for the interaction of the H_2 and H_3 molecules.

Since the relative stability of H_5Cl with respect to decomposition into other stable phases is of the order of few tens of meV, it is essential to ensure that temperature dependent contributions do not destabilize the *Cc* phase of H_5Cl. We expect indeed a temperature far from zero in typical experimental synthesis conditions under pressure where triatomic hydrogen would be produced. In view of that, we calculated the Gibbs free energy within the harmonic approximation at a temperature of 300 K. The Gibbs free energy curves as a function of pressure and of the fractional concentration of H are shown in Figure 2.20 [98], in comparison with the ones at 0 K.

Including the thermal effect to the phase diagram enhances the stability of H_5Cl as this composition becomes thermodynamically stable at a lower pressure, even below 50 GPa. Furthermore, the transition pressures between the stable phases slightly shift, as indicated by the bars shown on the basal plane of Figure 2.20 [98]. An additional phase is stabilized between 145 and 185 GPa for the stoichiometry H_3Cl, which contains HCl zig-zag chains, H_2 and, again, H_3 molecules, showing that the system is in a frustrated state between diatomic and triatomic hydrogen. Eventually, at higher pressures, phases containing H_3 clearly dominate the phase diagram of H-rich stoichiometries.

The structural changes on the H_5Cl structure described before can be further analyzed with the help of Figure 2.21, which shows the electron density at different pressures. First, at 0 GPa H_5Cl is unstable: HCl and H_2 molecules are separate. At 50 GPa, the structure is very close to the convex hull (see Figure 2.16): the H-Cl bond stretches to 1.37 Å while the H atom (which is rather just a proton, as most of the charge is kept around the more electronegative Cl atom) gets closer to the H_2 dimer an starts to form a triangle. This effect can be seen better at 100 GPa. Calculated Bader charges shows that in the studied range of pressures the charges on the Cl atoms remain nearly unchanged (-0.40∼-0.5), hence the H_2 dimer shares its charge with a proton to form a H_3^+ molecule. We can conclude that the H_3 units can be seen as stable H_3^+ ions immersed in a background of negative charge.

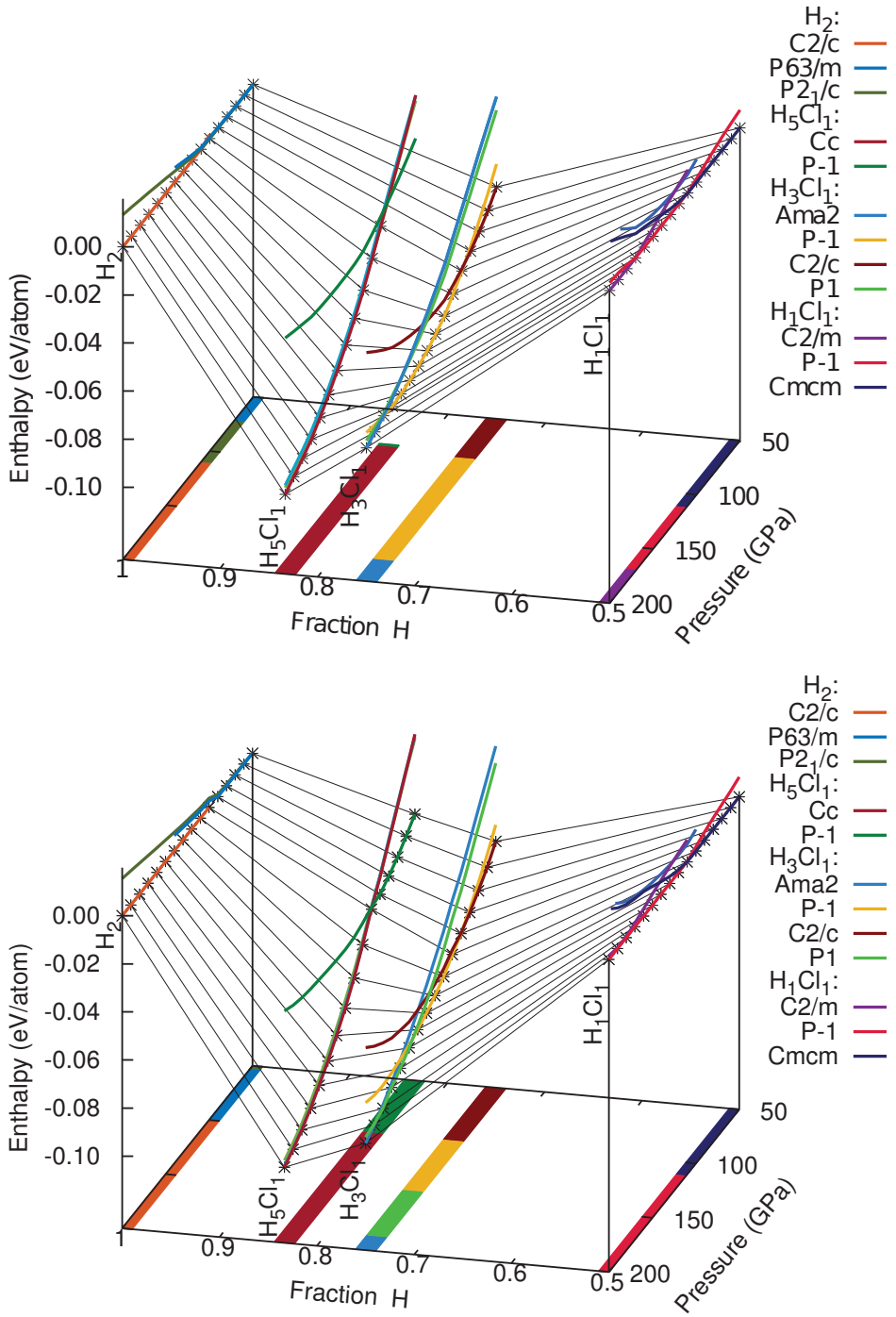
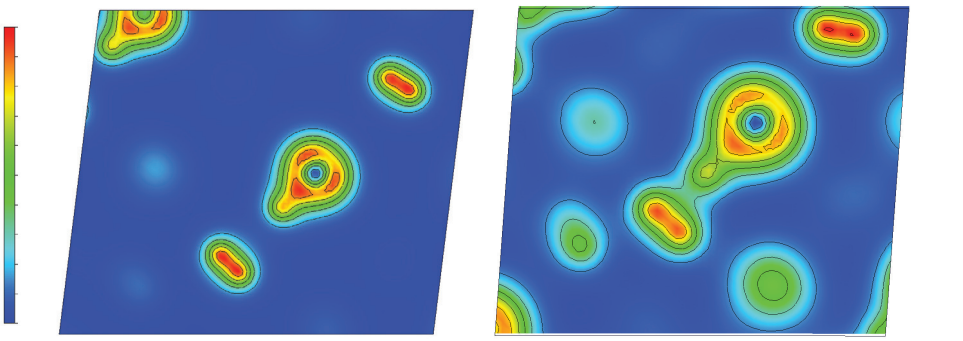
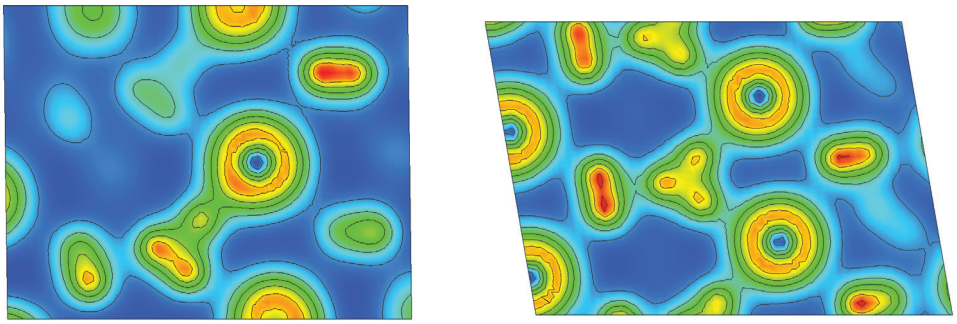


Fig. 2.20 Top Panel: Enthalpies per atom (static lattice) in the hydrogen rich regime of the compositional space as a function of pressure and fractional concentration of H for all competing low-lying phases. Stable phases are marked with black crosses and projected onto the basal plane. Bottom Panel: Free energies per atom in the harmonic approximation at 300 K for the same phases that are shown in the left panel. Note that, at finite temperature, the composition H_5Cl_1 becomes thermodynamically stable at a pressure below 50 GPa.



(a) 0 GPa (b) 50 GPa



(c) 100 GPa (d) 150 GPa

Fig. 2.21 Slices of the electronic density in the H_5Cl cell at different pressures.

It is important to remark that H_3 triangular clusters were predicted to become stable in pure hydrogen above 2.1 TPa [99]. Therefore, triangular H_3 appears in H-Cl compounds at pressures 20 times smaller than those necessary to make these exotic structures stable in pure hydrogen. This unexpected finding is extremely thrilling as pressures of 100 GPa are easily accessible in experimental setups, while few TPa are at present impossible to reach in static conditions on earth. The possibility to obtain a stabilized H_3 molecule in H_5Cl at moderate pressures would enable the direct experimental study of properties of triatomic hydrogen, obtaining spectra to be compared with theoretical calculations and astronomical detection of this extremely important and elusive ion. Furthermore, the synthesis of these predicted non-conventional stoichiometries of HCl under moderate pressures could considerably expand our knowledge of chemical bonds in hydrogen under pressure.

Chapter 3

Materials for Photovoltaics

This chapter is devoted to the study of materials with applications in photovoltaics: the transparent conductive oxide $CuBO_2$, novel nitride perovskites and $Cu(In,Ga)S_2$ chalcopyrite alloys. We focus here in the importance to determine the crystal structure to investigate the interplay between structural and electronic properties. First, we prove from global structure prediction, that the delafossite structure of $CuBO_2$ proposed in 2007 by Snure and Tiwari [100], is very unlikely the ground state of this material. In fact, we found several other crystal structures lower in energy by up to 600 meV/atom. In the second section we present our results on the possibility of the existence of nitrogen perovskites. Our approach is in this case, a combination of high-throughput techniques and global structure prediction. We find 21 new compositions of the form ABN_3 that are thermodynamically stable and that have therefore excellent chances of being experimentally accessible. Most of these materials crystallize in monoclinic phases, but three compounds, namely $LaReN_3$, $LaWN_3$ and $YReN_3$, are predicted to have distorted perovskite structures as their ground state. In the last section we discuss the study of the element-specific electronic states of $Cu(In,Ga)S_2$ chalcopyrite as a function of the In/Ga ratio. We performed Density Functional Theory calculations to understand why X-ray absorption experiments observed a shift of the K and L_3 S absorption edges with the change of the In/Ga ratio, while the Cu, Ga, In absorption edges positions remained constant. Our calculations indicate that this behavior results from the dependence of the electronic states on the local atomic environment.

3.1 Crystal Structure of the P-type Transparent Conductive Oxide $CuBO_2$

We employed the Minima Hopping to obtain the ground-state structure of $CuBO_2$. This is a very promising p-type transparent conductive oxide that was recently synthesized, and it was thought to belong to the delafossite family. We proved that the true ground state is not the delafossite structure, and that the most promising candidate is a low symmetry monoclinic phase. This is still a layered structure, but with boron and copper

having a different coordination with respect to the delafossite phase.

Transparent conducting oxides (TCOs) are a class of semiconductors which are simultaneously rather good conductors of electricity and transparent to the visible frequencies of the electromagnetic spectrum. They posses carrier concentrations of at least 10^{20} cm^{-3} (to be compared with other pure semiconductors $10^{6} - 10^{13}$ cm^{-3}) and optical band gaps greater than 3 eV [101, 102]. TCOs have an important role in devices requiring a transparent contact like solar cells, LCD displays and low thermal emittance coatings [103, 104].

These materials offer even much more than transparent contacts, namely, the fabrication of transparent transistors for development of transparent electronics. The major ingredient missing for large-scale development of such technologies is a p-type semiconductor with a large gap, high conductivities and mobilities, together with a controlled transparency and that can be manufactured industrially. High performance n-type TCOs such as SnO_2, In_2O_3 and ZnO are part of our everyday life, however, p-type TCOs have been harder to develop [103]. The most promising materials are still probably Cu-based delafossites, such as $CuAlO_2$, the first report of a p-type TCO, by Kawazoe *et al.* in 1997 [105]. Since then, other $CuXO_2$ delafossites that combine transparency and p-type conductivity have been synthesized, namely X = Sc, Y, In, Ga, and Cr.

Among all these Cu-based compounds, $CuBO_2$ is particularly interesting as it possesses a large band gap (direct and indirect) and higher intrinsic conductivity than any other Cu-based TCO [100]. Furthermore, it has a crystal structure compatible with the one of α-Al_2O_3 sapphire and ZnO, the latter being already widely used as n-type TCO. This last point is important for the design of a transparent p-n junction, the fundamental brick of any active electronic device.

$CuBO_2$ thin films were for the first time synthesized and characterized in 2007 by Snure and Tiwari [100]. However, a theoretical work by Scanlon *et al.* [106] questioned the validity of the reported experimental lattice parameters, as they were in disagreement by more than 10% with the theoretical values obtained using either the local density approximation or the Heyd–Scuseria–Ernzerhof (HSE06) [81, 82] hybrid functional. This result was unexpected, as Density Functional Theory (DFT) calculations usually (and, in particular, for the other Cu delafossite structures) give good estimations of the lattice parameters. Moreover, the direct band gap calculated with accurate many-body methods for the $CuBO_2$ delafossite structure relaxed within the local density approximation yielded 3.52 eV [107], much smaller than the reported

Table 3.1 Phases of $CuBO_2$ found in the MH calculations.

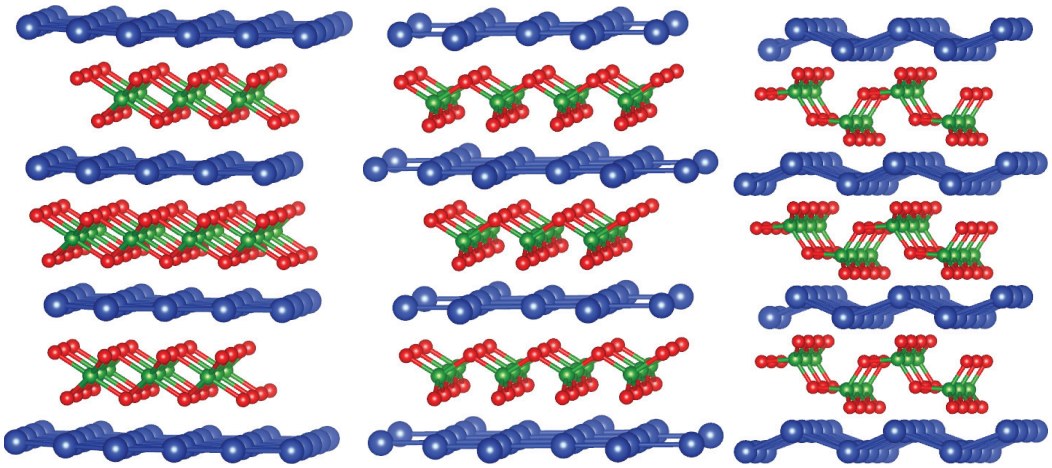
Space group	Energy (PBE) (eV/atom)
Cc	-6.751
$I\bar{4}m2$	-6.665
$Cmc2_1$	-6.664
$R\bar{3}m$	-6.183
$P6_3/mmc$	-6.182

experimental optical gap of 4.5 eV [100]. On the other hand, the same technique yielded, for other delafossites ($CuAlO_2$, $CuGaO_2$, and $CuInO_2$), direct band gaps consistently larger than experiment [107] (with the difference being due to excitonic effects and lattice contribution to the screening [108]). These facts were very surprising, and led to identifying $CuBO_2$ as an atypical compound within the class of Cu delafossites.

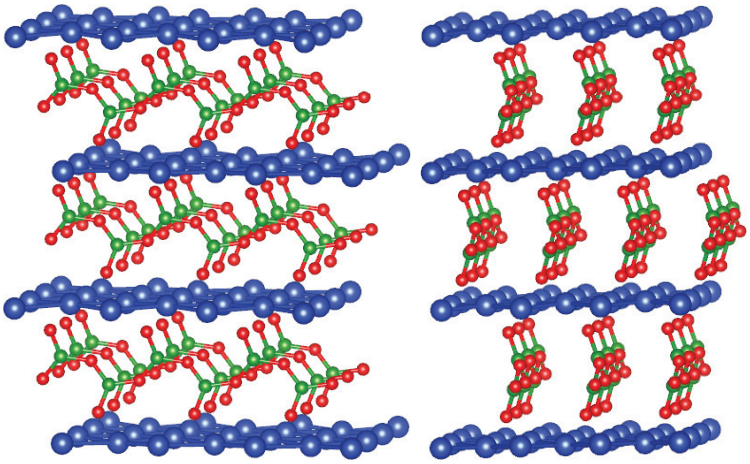
To shed more light on this controversial issue, we performed Minima Hopping (MH) calculations of $CuBO_2$ cells containing 1, 2, and 3 formula units (up to 12 atoms). Forces and energies were obtained using DFT as implemented in the code VASP [37, 38] with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [12]. In the extensive structure search, no low-energy local minimum with lattice constants similar to the ones given by Snure and Tiwari (a=2.84 Å and c=16.52 Å) [100] was found.

The most stable structure turned out to be a low-symmetry atomic arrangement, that the software ISOTROPY [109] classified as belonging to space group Cc with $a = 3.63$ Å, $b = 12.45$ Å, $c = 4.42$ Å, $\alpha = \gamma = 90°$, and $\beta = 86.78°$, with Cu atoms at the 4a Wyckoff position (-0.0055, 0.2465, 0.0439), B at the 4a Wyckoff position (-0.3632, 0.4496, 0.1981), and two non-equivalent O at the 4a Wyckoff positions (0.2555, 0.1490, -0.2105) and (0.0791, 0.0349, 0.3937). This crystal structure, which can be seen in Figure 3.1d, has planar Cu layers alternating with linear BO_2 networks where B is threefold coordinated. Delafossite structures with the two possible stacking configurations of the layers (hexagonal (space group $P63/mmc$) and rhombohedral (space group $R\bar{3}m$) [110]), appear in our simulations. However, they were not, by far, the lowest energy phases. Other interesting high-symmetry phases found during our MH runs are listed in Table 3.1 and shown in Figure 3.1.

The Cc structure lies almost 0.6 eV per atom lower in energy than the $R\bar{3}m$ delafossite, followed about 0.1 eV per atom higher by the $Cmc2_1$ and $I\bar{4}m2$ phases, which are almost degenerate. Between the ground state and the delafossite structure, many other phases (mostly with very low symmetry) were found. Cross-check MH



(a) $R\bar{3}m$ (b) $I\bar{4}m2$ (c) $Cmc2_1$



(d) Cc (a) (e) Cc (c)

Fig. 3.1 Schematic representation of the $CuBO_2$ structures: (a) delafossite, (b) tetragonal ($I\bar{4}m2$) (c) orthorhombic $Cmc2_1$, (d) and (f) Monoclinic Cc along the a and c crystallographic directions

runs for $CuAlO_2$, $CuGaO_2$, and $CuInO_2$ yielded invariably the delafossite as the lowest energy structure. Furthermore, the lower symmetry Cc and $Cmc2_1$ structures are not stable for any of those systems.

Although an energy difference of 0.6 eV is considerably larger than the typical error incurred in PBE calculations, we performed calculations (using the PBE geometry) at higher levels of theory and also for the other delafossites to confirm our findings. We used the HSE06 screened hybrid [81, 82], the PBE0 hybrid [111], and GGA + U as described in Reference [112] setting U−J = 5 eV for the d-states of Cu. Our results are

Table 3.2 Energy per atom (in meV) of different structures of $CuXO_2$ studied in this work using different approximations. The zero is the energy of the delafossite $R\bar{3}m$ structure.

X	B Cc	B $Cmc2_1$	B $I\bar{4}m2$	Al $I\bar{4}m2$	Ga $I\bar{4}m2$	In $I\bar{4}m2$
PBE	-569	-483	-482	123	138	248
PBE+U	-598	-508	-491	113	128	236
HSE06	-660	-565	-555	124	135	262
PBE0	-659	-546	-546	131	143	289

summarized in Table 3.2. Results for the $P6_3/mmc$ phase were basically identical (to a few meV) to those for the $R\bar{3}m$ structure and are therefore not shown in the table.

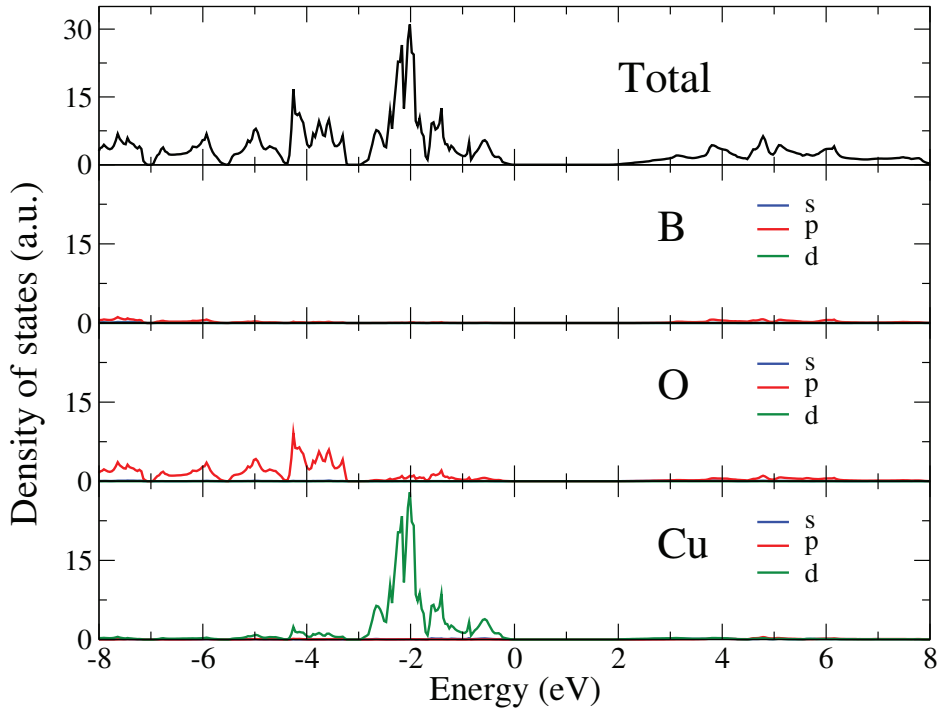


Fig. 3.2 Partial density of states of Cc $CuBO_2$ calculated with PBE+U. The Fermi level is shifted to zero.

The four approximations (PBE, PBE+U, HSE06, and PBE0) yield essentially the same energy differences for all structures and compounds studied ($CuBO_2$, $CuAlO_2$, $CuGaO_2$, and $CuInO_2$). Furthermore, and confirming our MH simulations, the Cc structure is only the ground state for $CuBO_2$.

A possible explanation for the destabilization of the delafossite phase of $CuBO_2$ can be found by analyzing the oxidation state of the metal. In fact, in the delafossite structure, the metal atom is in the oxidation state +3. However, by performing a Bader analysis of the delafossite $CuBO_2$ one finds a Bader charge on the B atom of +2.15, to compare with +2.54 for Al in $CuAlO_2$. This charge increases to +2.31 in the *Cc* $CuBO_2$

All considered phases of $CuBO_2$ turned out to have indirect fundamental band gaps, and all band gaps are large enough to assure transparency in the visible. The top of the valence band displays a mixed Cu *d* with O *d* character while the bottom of the conduction is a mixture of Cu *s*, *p*, and *d* states with mainly O *p* (and some O *s*) states, and a significant amount of B *p*. This is similar to the case of the delafossite structure [107]. Moreover, we observed that the projected densities of states of different phases do not show qualitatively different features.

In conclusion, our results indicate that it is extremely unlikely that the experimentally produced phase belongs to the delafossite family. The most likely candidate is instead a low symmetry monoclinic structure, consisting of alternating planar Cu layers (fourfold coordinated) and BO_2 networks (where B is threefold coordinated). These results call for further experimental work to determine unambiguously the structure of this promising p-type TCO.

3.2 Prediction of Stable Nitride Perovskites

Perovskites are one of the most studied classes of materials, with a variety of applications in diverse fields of science and technology. Their basic composition is ABX_3, where X is a non-metal normally from the VIA or VIIA group. We investigate the possibility of the existence of perovskites with X=N. Our approach is based on a combination of high-throughput techniques and global structure prediction methods. We find 21 new compositions of the form ABN_3 that are thermodynamically stable (considering all possible decomposition channels) and that have therefore excellent chances of being experimentally accessible. Most of these materials crystallize in monoclinic phases, but three compounds, namely $LaReN_3$, $LaWN_3$ and $YReN_3$ are predicted to have distorted perovskite structures in their ground state. In particular, $LaWN_3$ is a semiconductor and displays a large ferroelectric polarization.

The perovskite structure is one of the most common and most extensively studied in materials science. The general chemical formula for perovskite compounds is ABX_3, where A and B are two cations of different sizes, and X is an anion that binds to both.

Its ideal structure is cubic, with the B atoms in the center of a characteristic octahedra of anions. This seemingly simple atomic arrangement hides an exceptional diversity of physical and chemical properties.

In fact, perovskites exhibit a wealth of remarkable properties that lead to applications in numerous technological fields. A prominent example are piezoelectric perovskites, such as lead zirconate titanate, where the piezoelectric effect is used for sensor or actuator applications [113]. Another example are perovskite high temperature superconductors, such as yttrium barium copper oxide [114, 115]. Some materials of the perovskite family, mostly manganese-based perovskite oxides, exhibit colossal magnetoresistance that enables them to dramatically change their electrical resistance in the presence of a magnetic field [116]. Also many multiferroic materials, i.e., phases that support the co-existence of more than one primary ferroic order parameter simultaneously, are transition metal oxides with the perovskite structure [117, 118]. More recently, there was a surge of interest in halide perovskites due to their use as absorbing materials for solar cells [119, 120]. Moreover, perovskites are studied in several other fields: as possible thermoelectric materials [121], as catalysts [122, 123], as lasers [124, 125], etc. Finally, we remark that silicate perovskites are one of the most common minerals in the lower part of Earth's mantle.

There are two main reasons for such great variety of physical and chemical applications. First, the perovskite structure can accommodate a large number of possible A and B elements, spanning a large portion of the periodic table. Second, the cubic ideal structure can be distorted in several different ways, e.g. by tilting of the octahedra or by displacing the cations from the center of their polyhedra. We focus on the first aspect: the unparalleled chemical variety of the perovskites.

The best known, and perhaps most common, perovskites are oxides, i.e., the anion X is O^{2-}. In this case, the cations can have oxidation states $A^{+}B^{5+}$, $A^{2+}B^{4+}$, or $A^{3+}B^{3+}$. A smaller number of fluorides, sulfides, other halides, and a few selenides are also known to crystallize in the perovskite structure. Other more complex phases also exist. For example, oxynitrides (of the type ABO_2N) with perovskite structures are an emerging class of materials due to their interesting optical, photocatalytic, dielectric and magnetoresistive properties [126]. In this case, the nitrogen and the oxygen atoms may occupy the same sites, forming solid solutions with ratios that can be adjusted to some extent during the process of synthesis.

We have taken a step further and investigate if it is thermodynamically possible to create nitride perovskites, i.e. with compositions of the type ABN_3. Assuming the standard oxidation state of -3 for nitrogen, this would imply very high oxidation

states for the cations of at least +5 to +7. Although several transition metals support very high oxidation states, however nitrogen is less electronegative than oxygen and we can expect that such high oxidation states are less likely to be achieved. It is therefore reasonable to expect that the nitrogen perovskite structure can only be stable in a very narrow domain of composition space. On the other hand, for the same reason we can anticipate unusual electronic properties. Above all, the existence of a completely new class of ABN_3 perovskites with potential interesting properties would once more prove the modernity of this 175-year "old" structure.

Unfortunately, it is not possible with current computer resources to investigate a large number of ABN_3 compounds using global structure prediction. We remind that in Chapter 2, we calculated 33 stoichiometries for each binary phase diagram, determined by considering structures with up to 12 or 13 atoms per cell. When it comes to perovskites, not only there is a large number of possible ABN_3 stoichiometries, but also they often crystallize in deformed structures with a relatively large number of atoms in the unit cell (10–40). We chose therefore to employ a mixture of prototype search and global structure prediction with the Minima Hopping (MH) method [25, 26] that proved to be very efficient for our problem.

All calculation were performed at zero temperature and pressure, and the effects of the zero point motion of the phonons (that are not expected to be important here) were neglected. In this case, the thermodynamic quantity of interest is the total energy. It is clear that to calculate the distance to the convex hull for our systems requires the evaluation of the total energy of all binary and ternary phases that contain nitrogen. This is certainly an essential step but also tedious and time-consuming. Fortunately, there are by now a number of excellent freely available databases, like the Materials Project [127], the AFLOWLIB [16], or the Open Quantum Materials Database [17] that provide this information (among a wealth of other data). In the following we will use the materials project database for our reference energies, and the distance to the convex hull is calculated with the help of PYMATGEN [128].

The energies and forces were calculated with Density Functional Theory within the projector augmented wave (PAW) method as implemented in the code VASP [37, 38]. We used the PAW datasets of version 5.2 (for compatibility with Materials Project [127] and Open Quantum Materials Database [17]). A cutoff of 520 eV and dense k-point grids were selected to ensure an accuracy of 2 meV/atom in the total energy. All forces were converged to better than 0.005 eV/Å. To approximate the exchange-correlation functional of density functional theory we used the Perdew-Burke-Ernzerhof [12] generalized gradient approximation.

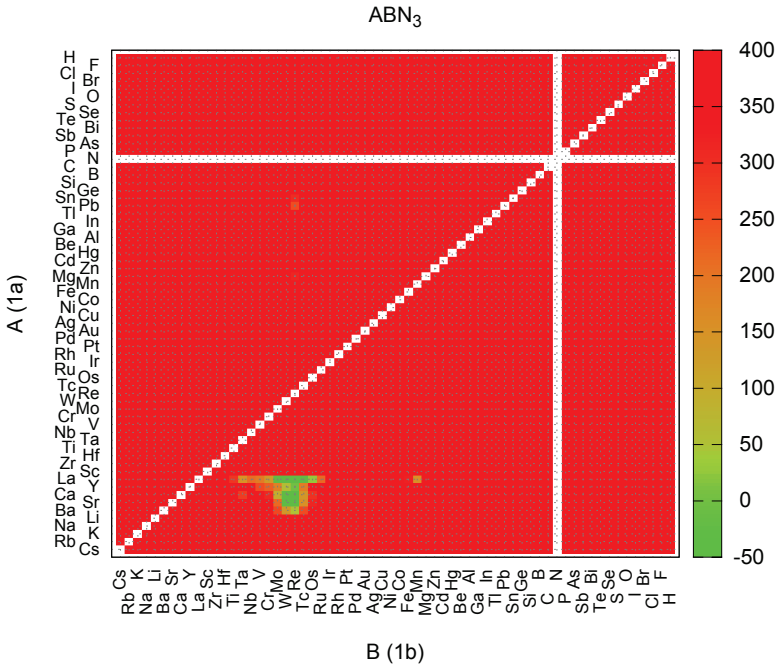


Fig. 3.3 Distance to the convex hull of stability for the ABN_3 compositions in the cubic perovskite phase. The A atom is in the Wyckoff position $1a$ and the B atom in the $1b$. To order the atoms along both axes use the modified Pettifor ordering of Ref. [129].

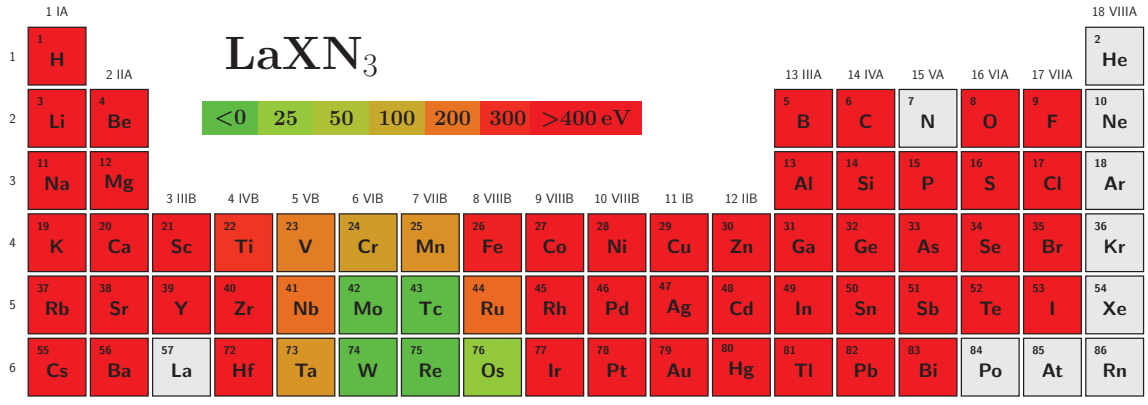


Fig. 3.4 Distance to the convex hull of stability for the $LaXN_3$ compositions in the cubic perovskite phase, where X runs over the periodic table. The color scale is identical to Figure 3.3.

Our starting point is the simple cubic perovskite phase with 5 atoms in the unit cell. We constructed crystal structures using this prototype for all combinations ABN_3, where A and B are all elements up to bismuth, but excluding rare gases (which do not form hard compounds) and lanthanides. We note that this choice was somewhat arbitrary, and ultimately decided by efficiency reasons. In the following we will use the notation that element A rests in the $1a$ Wyckoff position of the cubic $Pm\bar{3}m$ lattice and atom B in the $1b$ position. The lattice constant of each phase was then optimized using VASP (note that all atomic positions are determined by symmetry). We then obtained the distance to the convex hull of thermodynamic stability using PYMATGEN [128] and the Materials Project database [127].

The results of our high-throughput search are summarized in Figure 3.3, where we plot the distance to the hull for the 1953 stoichiometries (leading to 3906 structures) studied here. The elements are ordered using a modified [129] Pettifor scale [130], built in order to put into evidence the chemical similarity between the elements. Neighbor elements are to some extent similar in the sense that they often lead to similar crystal structures. The picture uses a color scale to indicate the distance to the convex hull: green for structures that are thermodynamically stable, that gradually becomes red for phases that are higher than 400 meV/atom. Note that we measure the distance to the hull defined solely by the structures present in the materials project database, so our distance will be negative for novel stable phases.

The first conclusion that we can read from the figure is that most possible combinations are highly unstable (with much more than 1 eV/atom higher than the convex hull). This is perhaps not surprising. Normal perovskites ABX_3 are quite ionic materials, with X being a rather electronegative atom from the oxygen group or an halogen. Let us take the example of an halogen, with its nominal charge of -1. Then the combined charge of the A and B atoms should sum to $+3$. Our systems contain nitrogen. Its nominal charge is -3, which would require that the combined A and B charge should be $+9$. Such large charges are certainly possible, but the number of possible A and B are quite reduced. Furthermore, nitrogen is considerably less electronegative than oxygen or fluorine (Pauling electronegativity of 3.0 in comparison with 3.5 for O, or 4.0 for F), which brings us farther from the optimal chemical conditions for the formation of the perovskite phase.

It is therefore quite surprising that, in spite of being far away from the sweet spot for stability, we do find that the cubic perovskite phase is thermodynamically stable in a relatively narrow region of composition space. The fact that stable compounds form a localized (green) region in Figure 3.3 is expected due to the used of the (modified)

Table 3.3 Properties of the new stable compounds: energy distance to the known convex hull of the cubic perovskite (CP) phase ΔE_{CP} in meV/atom; energy distance to the known convex hull of the ground-state (GS) phase ΔE_{GS} in meV/atom; space group number (spg); indirect (I) and direct (D) electronic band gap $E_{\mathrm{g}}^{I,D}$ calculated using the HSE06 hybrid functional (in eV) for the ground-state structure; Bader charges (BC) for element A, B, and for nitrogen.

	ΔE_{CP}	ΔE_{GS}	spg	E_{g}^{I}	E_{g}^{D}	BC (A)	BC (B)	-BC (N)	comment
$LaReN_3$	-356	-365	62	-	-	1.93	2.14	1.35	metallic perovskite
$LaWN_3$	-293	-353	161	1.98	2.23	1.93	2.45	1.46	semiconducting perovskite
$LaMoN_3$	-176	-295	15	2.84	3.00	1.95, 1.97	2.11	1.28, 1.38, 1.41	
$LaTcN_3$	-171	-224	15	-	-	1.93, 1.96	1.72	1.19, 1.21, 1.27	metal
$SrWN_3$	-164	-282	1	1.84	2.06	1.52	2.30, 2.41	0.98–1.45	
$CaWN_3$	-100	-277	12	2.21	2.28	1.52	2.37	0.92–1.50	
$CaReN_3$	-82	-280	15	2.49	2.71	1.54	2.17	1.13, 1.28, 1.30	
$SrReN_3$	-65	-225	15	2.71	2.84	1.55	2.14	1.14, 1.27, 1.30	
$YReN_3$	-2	-152	62	-	-	2.03	2.09	1.38	metallic perovskite
$LaOsN_3$	24	-22	51	-	-	1.82	0.92	1.35, 0.76, 0.63	metal
$BaReN_3$	50	-189	15	2.46	2.78	1.54, 1.57	2.13	1.13, 1.24, 1.31	
YWN_3	56	-191	15	3.10	3.32	2.07	2.41	1.41, 1.47, 1.59	
$CaMoN_3$	92	-179	12	1.58	1.58	1.51	2.09	0.80–1.40	
$LaCrN_3$	130	-81	15	2.13	2.17	1.95, 1.98	1.71	1.12, 1.26, 1.29	
$BaWN_3$	133	-52	14	2.08	2.11	1.50	2.11	0.82,0.86, 1.33,1.44	
$CaTcN_3$	139	-129	12	1.52	1.59	1.50	1.64	0.65, 1.19, 1.28	
$LaTaN_3$	146	35	14	1.27	1.37	1.87	2.33	1.08, 1.53, 1.59	
$SrTcN_3$	152	-68	15	2.33	2.46	1.55, 1.58	1.86	1.03, 1.17, 1.22	
$LaMnN_3$	153	-158	12	-	-	1.81	1.15	0.70, 0.80, 1.45	magnetic metal
$YMoN_3$	179	-136	15	2.69	2.95	2.07	2.15	1.31, 1.38, 1.53	
$YTcN_3$	186	-53	15	-	-	1.76	2.03	1.19, 1.21, 1.41	metal
$LaVN_3$	187	6	1	1.38	1.80	1.91, 1.89	1.73, 1.83	0.90, 0.96, 1.31–1.40	
$LaCoN_3$	508	-156	12	1.32	1.76	1.81	0.58	0.22, 0.68, 1.50	

Pettifor scale where neighboring atoms are chemically similar, and gives us further confidence in our results.

There are only 5 elements that can occupy the Wyckoff 1*a* position and yield a nitride perovskite, namely Ca, Sr, and Ba (group IIA) and Y and La (group IIIB). We plot, in Figure 3.4 the distance to the convex hull for $LaXN_3$, where X runs over the periodic table. It is clear that all elements that can be positioned in the Wyckoff position 1*b* of the perovskite structure are transition metals, namely Re, W, and several of their neighbors in the periodic table. This is of course the region of the periodic table where the highest oxidation states can be found.

These results prove that there are several compositions that are thermodynamically stable, and that are still unknown experimentally. This is already remarkable, as there are only two entries with compositions of the type ABN_3 in the Materials Project database [127], namely $BaClN_3$ and $SiPN_3$ (none of them a perovskite). From Figure 3.3 and Figure 3.4 we already see that the number of stable compositions of this family is likely an order of magnitude larger than what is currently known. We emphasize that for many of these compositions the distance to the convex hull is considerably larger than the theoretical error in this kind of calculations, so we are confident that the prediction of stability will very likely stand against experimental verification.

Up to now we have only been able to prove that there are a number of compositions that are thermodynamically stable, where the octahedral arrangement of the nitrogen atoms typical of the perovskite phase seems to be chemically favored. Moreover, it is well known that ABX_3 compounds rarely crystallize in the cubic perovskite structures, but prefer distorted phases with a larger number of atoms in the unit cell. For example, a quick glance at structures for ABX_3 compounds, (with X = N, O, S, Se, Te, F, Cl, Br, I), present in the materials project database shows that that only 110 materials appear with the cubic perovskite phase, while more than 350 have 10 atoms in the unit cell, and around 500 have 20. For these reasons, from a simple prototype search we can hardly provide any further statement concerning the ground-state crystal phase for these compositions.

To go beyond this situation we decided to use a mixed approach. We started to select all compositions for which the cubic perovskite phase is already either stable or within 200 meV per atom from the convex hull. This choice restricted the total number of compositions from almost 2000 to a mere 22. For these we performed structure prediction simulations using the MH method with two formula units (10 atoms) in the unit cell. At least 2 independent simulations were performed for each composition, and

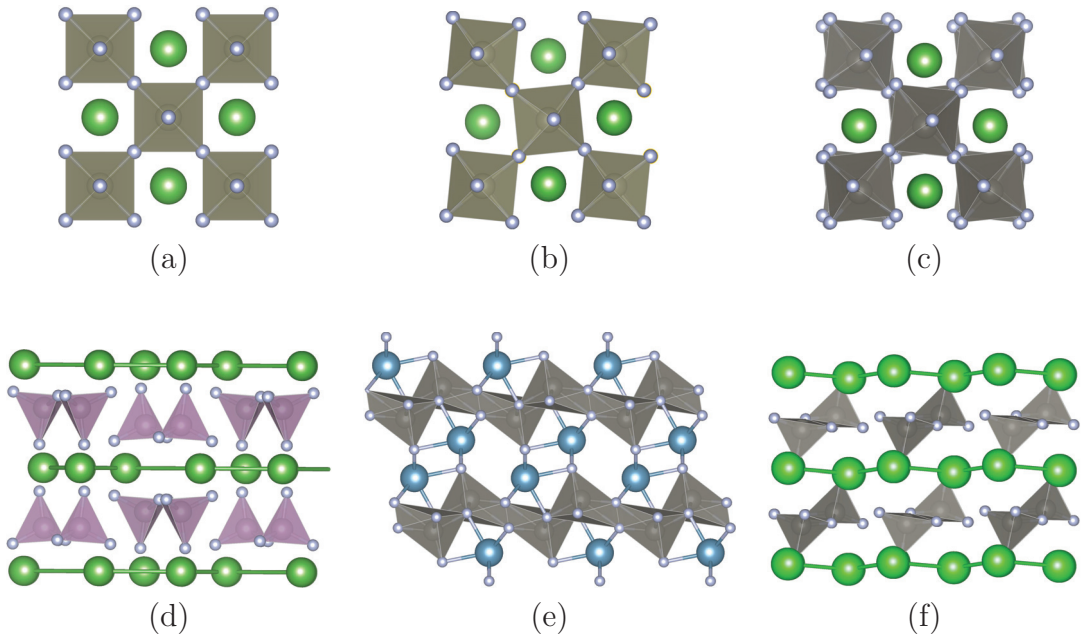
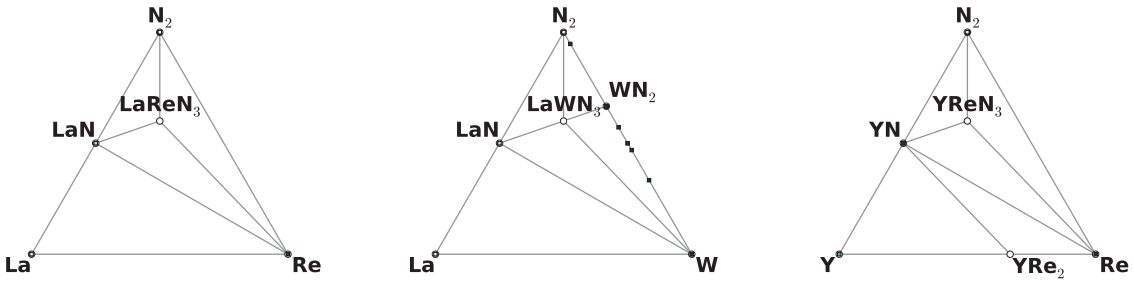


Fig. 3.5 The most relevant crystal structures obtained in our work: (a) the starting cubic perovskite structure (space group 221); (b) the orthorhombic perovskite structure (space group 62) of $LaReN_3$ and $YReN_3$; (c) the trigonal structure (space group 161) of $LaWN_3$; (d) The monoclinic structure (space group 15) adopted by $LaMoN_3$, $LaTcN_3$, $CaReN_3$, $SrReN_3$, $BaReN_3$, YWN_3, $LaCrN_3$, $SrTcN_3$, and $YTcN_3$; (e) The monoclinic structure (space group 12) adopted by $CaWN_3$, $CaMoN_3$, and $CaTcN_3$; (f) The monoclinic structure (space group 14) adopted by $BaWN_3$ and $LaTaN_3$.

each simulation detected at least 50 minima. In the cases where the simulations were inconclusive further MH runs were performed.

Unfortunately, structure prediction runs with more than 10 atoms per units cell turned out to be computationally prohibitive for these systems. We decided therefore to complement this technique by also testing the most common phases in which ABX_3 compounds crystallize. We used a total of 16 prototypes with 5–20 atoms in the unit cell. This combined approach gives us a rather complete picture of the energy landscape of our compounds, and allows us to have a large degree of confidence in our ground-state structures for the large majority of systems.

Our results are summarized in Table 3.3 where we list the distance to the hull of the cubic perovskite phase, the distance to the hull of the ground-state structure, and the space group, the electronic indirect and direct band gap computed with the Heyd-Scuseria-Ernzerhof approximation (HSE06) [81, 82] and the Bader charges of the ground-state structure. The electronic band gaps were obtained with the HSE06 hybrid functional to circumvent the severe underestimation typical of the PBE approximation.



(a) (b) (c)

Fig. 3.6 Theoretical ternary phase diagrams for (a) $LaReN_3$, (b) $LaWN_3$, and (c) $YReN_3$.

We can see that we now have 20 compositions that are below the convex hull, and are therefore potentially stable. As expected, none of these has the simple cubic perovskite crystal structure, as we always find a structure (either from the MHM simulations or from other prototypes) whose energy is lower. However, there are a few that do crystallize in a distorted perovskite structure, namely $LaReN_3$, $YReN_3$, and $LaWN_3$. The first of these crystallizes in an orthorhombic (space group 62, see Figure 3.5b) that can be obtained by a deformation of a $\sqrt{2} \times \sqrt{2} \times 2$ supercell of the original cubic perovskite. The breaking of the symmetry for $LaReN_3$ is small, which can also be inferred by the small difference of total energy between the orthorhombic and the cubic phases (9 meV/atom). The same structure is adopted by $YReN_3$, although with a larger symmetry breaking (energy difference 150 meV). These two perovskite phases seem to be metallic. On the other hand, the breaking of the cubic symmetry in $LaWN_3$ leads to a semiconducting trigonal lattice (space group 161, see Figure 3.5c). From Table 3.3 we can see that the bonding is quite ionic in these three perovskites, with around two electrons leaving from the A and B atoms to go to the basin of attraction of the nitrogens (gaining a net charge of around -1.5 electrons each). The ternary phase diagram for these systems, calculated with PYMATGEN and the materials project database, can be found in Figure 3.6.

Note that for the two La compounds are quite below the known convex hull (by more than -300 meV/atom), which is much larger than the theoretical error we can expect for this kind of calculation. We can therefore state that it is extremely likely that perovskites of the type ABN_3 can be produced experimentally in thermodynamic conditions. Furthermore, it is well known that the lanthanides can be, to a large extent, substituted by each other. Therefore, it is reasonable to expect that lanthanide (or actinide) based nitride perovskites are also stable. This is particularly interesting as

the localized f states of the lanthanides often lead to interesting magnetic properties with important technological applications.

As we can see from Table 3.3, all other stoichiometries studied here crystallize in structures very different from the perovskite structure. The most likely structure is monoclinic with space group 15 (depicted in Figure 3.5d), where the nitrogen atoms are no longer arranged in octahedra, but in tetrahedra, and the A atoms form flat layers. This structure seems to be adopted by $LaMoN_3$, $LaTcN_3$, $CaReN_3$, $SrReN_3$, $BaReN_3$, YWN_3, $LaCrN_3$, $SrTcN_3$, and $YTcN_3$. Most of these materials are semiconducting, with the exception of $LaTcN_3$ and $YTcN_3$ that are metallic.

The compounds $CaWN_3$, $CaMoN_3$, and $CaTcN_3$ are semiconductors that crystallize in a monoclinic phase (space group 12, Figure 3.5e). The third most likely structure is also monoclinic (space group 14, see Figure 3.5f), and composed also of A layers separated by nitrogen tetrahedra. This structure leads to semiconducting phases, and is adopted by $BaWN_3$ and $LaTaN_3$.

We would like to note that these compositions do not clearly exhaust all possible thermodynamically stable compounds with composition ABN_3. In fact, we steered our investigation towards finding new perovskite structures. It is perfectly plausible to believe that many other materials of the type ABN_3 can exist, although with a structure considerably different from the typical arrangement of atoms in the perovskites. As an example, we decided to perform structure prediction runs for $LaCoN_3$. The cubic perovskite structure for this composition is more than 500 meV/atom above the convex hull of stability, much higher than our initial cutoff of 200 meV/atom. However, the MHM method discovered a monoclinic structure (space group 12) that is below the convex hull by 156 meV/atom (i.e., 664 meV/atom lower than the cubic phase). Of course the ground-state structure is quite different from a perovskite, and consists of a layered structure. This shows that probably many more thermodynamically stable ABN_3 materials are still waiting for our discovery.

We propose an efficient approach to search for new stable materials, that combines standard prototype search and state-of-the-art global structure prediction methods. This method was applied to the discovery of new stable ABN_3 perovskite structures. From the ∼2000 possible compositions of the form ABN_3 , only two stable materials that do not crystallize in the perovskite structure were known experimentally. With our approach we found 21 new compositions that have an energy below the known convex hull of thermodynamic stability. This, by itself, shows that modern computational techniques offer an effective tool to screen thousands of possible materials in a fraction of the time required by an equivalent experimental search.

From the 21 new stable stoichiometries, most crystallize in monoclinic lattices where the nitrogen atoms are arranged in pyramids. However, three of them, namely $LaReN_3$, $YReN_3$ and $LaWN_3$, are likely to favor the perovskite structure. In particular, $LaWN_3$ is a semiconductor and displays a large ferroelectric polarization. This unexpected finding of stable nitride perovskites puts into question the accepted domain of stability of this crystal structure and opens the way for the experimental synthesis of exotic perovskites with new potential applications.

3.3 The Cu(In,Ga)S_2 Chalcopyrite Alloys

Near Edge X-Ray Absorption Fine Structure measurements performed on $CuIn_xGa_{1-x}S_2$ thin films for x=0, 0.67 and 1, by our experimental collaborators [131], show that the S absorption edges shift with changing In/Ga ratio as expected from the variation of the band gap. On the other hand, the Ga, In, and Cu absorption edges positions remain nearly unaltered with alloy composition despite the significant change of the band gap. We performed a DFT study of the element-specific unoccupied electronic states as a function of the In/Ga ratio. Our calculations indicate that this behavior originates from the dependence of the electronic states on the local atomic environment while the change in band gap arises from a spatial average of these localized states with changing alloy composition.

Photovoltaic conversion efficiencies of more than 20% have been demonstrated on both glass substrates and flexible polymer foils for Cu(In,Ga)Se_2 (CIGSe) based thin film solar cells. The properties of this absorber material can be tuned not only by adjusting the In/Ga ratio and also by substituting Se for S, obtaining Cu(In,Ga)(S,Se)$_2$ (CIGSSe) alloys, which is particularly interesting from the industrial point of view. Regarding the sulfides (CIGS), the band gap changes from 1.5 eV for $CuInS_2$ (CIS) to 2.4 eV for $CuGaS_2$ (CGS) caused by a small shift of the valence band maximum and a much larger shift of the conduction band minimum.

$CuInS_2$ and $CuGaS_2$ crystallize in the chalcopyrite lattice structure [132]. This is a tetragonal cell with 8 atoms and space group $I\bar{4}2d$ number 122. It is a diamond like structure similar to the zincblende ZnS double unit cell, but with the Zn-sites substituted by Cu and Ga/In. The precise location of the S atoms in the chalcopyrite structure depends on the strength of Cu-S and Ga-S/In-S bonds. A stronger bonding of S atoms with the two nearest Cu atoms was suggested presumably due to the *d*-electrons of the Cu atom contribution to the bonding, leading to a difference of the

strengths of the Cu-S and Ga-S/In-S bonds and therefore of the Cu-S and Ga-S/In-S bond lengths.

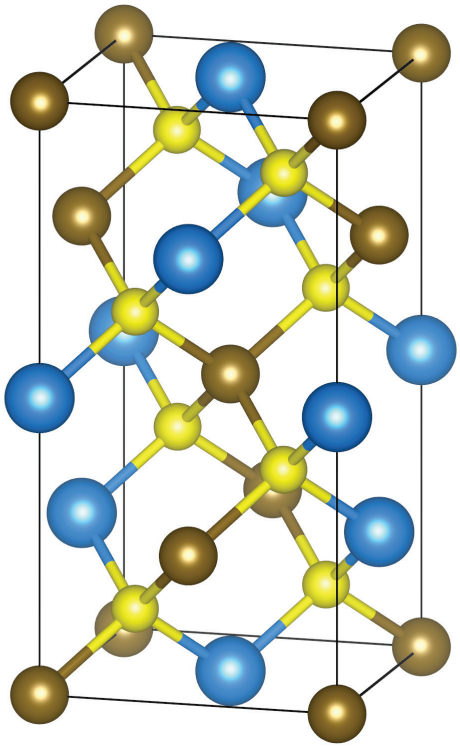


Fig. 3.7 Schematic representation of the atoms on the chalcopyrite structure. Cu: brown, Ge/In: blue and S: yellow. In the case of the double cell of zincblende, the Cu and Ga/In sites would correspond to the Zn sites.

The tetragonal distortion describing the new position of anions makes the lattice constant ratio c/a in the double zincblende structure, to deviate from 2 to ∼1.96 for $CuGaS_2$ and $CuInS_2$. The S anions are displaced from their ideal tetrahedral positions by an amount u known as the anion displacement parameter of the chalcopyrite structure. The parameter can be determined as

$$u = \frac{d^2_{Cu-S} - d^2_{Ga-S}}{a^2} - \frac{1}{2},$$

where d_{Ga-S} and d_{Ga-S} are the two anion-cation distances, and u is the value to add to the ideally tetrahedral anion coordinate leading (3/4+u,1/4,1/8) at Wyckoff possition 8d.

Unoccupied states can be accurately probed by Near Edge X-ray Absorption fine Structure (NEXAFS) spectroscopy. NEXAFS is an absorption-spectroscopy technique that allows to extract the features in the absorption spectra close the the absorption

edges. It involves the excitation of electrons from a core level to partially filled and empty states. The absorbed radiation is measured for a range of frequencies around the energy of the edges. It is element-specific as the elements have different core level energies. The K-edge for the Cu, Ga, In and S are 8978.9 eV, 10367.1 eV, 27939.9 eV and 2472.0 eV respectively. Therefore, the NEXAFS absorption signal contains information about the empty density of states of the material as well as transition probabilities.

Figure 3.8 shows the S, Ga and Cu K-edges and the In L_3-edge spectra. On the x-axis of the NEXAFS spectra we can read the excitation (photon) energy. The spectra from a single element in each figure have a common, although not necessarily absolute energy scale so that real shifts between the absorption edges can be considered. As expected, the S K-edge position in Figure 3.8a shifts to higher energies with the Ga content of the material and thus the band gap is increased. In contrast, no change in edge position is observed for the Ga or Cu K-edges nor for the In L_3-edges. This is also observed for Ga L_3-edges, and the In $M_{4,5}$-edge, while the S L_3-edge exhibits again a significant shift (see Figure D.1 [131]).

In order to explain this behavior, we performed DFT calculations of the electronic states of $CuIn_xGa_{1-x}S_2$ alloys, for x= 0, 0.3, 0.5, 0.67, and 1. The projected partial densities of states (pDOS) exhibit the same dependence on or independence of the composition as the measured NEXAFS spectra.

First, in order to obtain a microscopic model of the alloy, we used special quasirandom structures (SQS) [133] to approximate random arrangements by using small periodic supercells. SQSs are periodic supercells with a small number of atoms (starting from N=8), that mimic at best the first few, physically most relevant, radial correlation functions of a perfectly random structure. The idea is to design a special N-atom structure whose distinct (i.e. not ensemble average) correlation functions best matches the ensemble average of the random alloys. The electronic states of the CIGS compounds are known to be very sensitive to modifications of the internal structural parameters. This fact necessitates a proper description of the microscopic atomic structure of CIGS alloys, and it motivated the choice to work with supercells, despite the increased computational cost.

The SQS of $CuIn_xGa_{1-x}S_2$ for $x \simeq 1/3$, $x \simeq 1/2$, and $x \simeq 2/3$ were built using the ATAT [134] code. The SQS supercells contained 40, 54, or 64 atoms and they include all three possible environments for S: 2 Cu and 2 Ga neighbors, 2 Cu and 2 In neighbors, and 2 Cu, 1 Ga, and 1 In neighbors as described in Reference [132]. No defects were considered (e.g., cation antisites) so that all S atoms are surrounded by 2 Cu and 2 atoms of the group IIIA, yielding always the same number of valence electrons. The

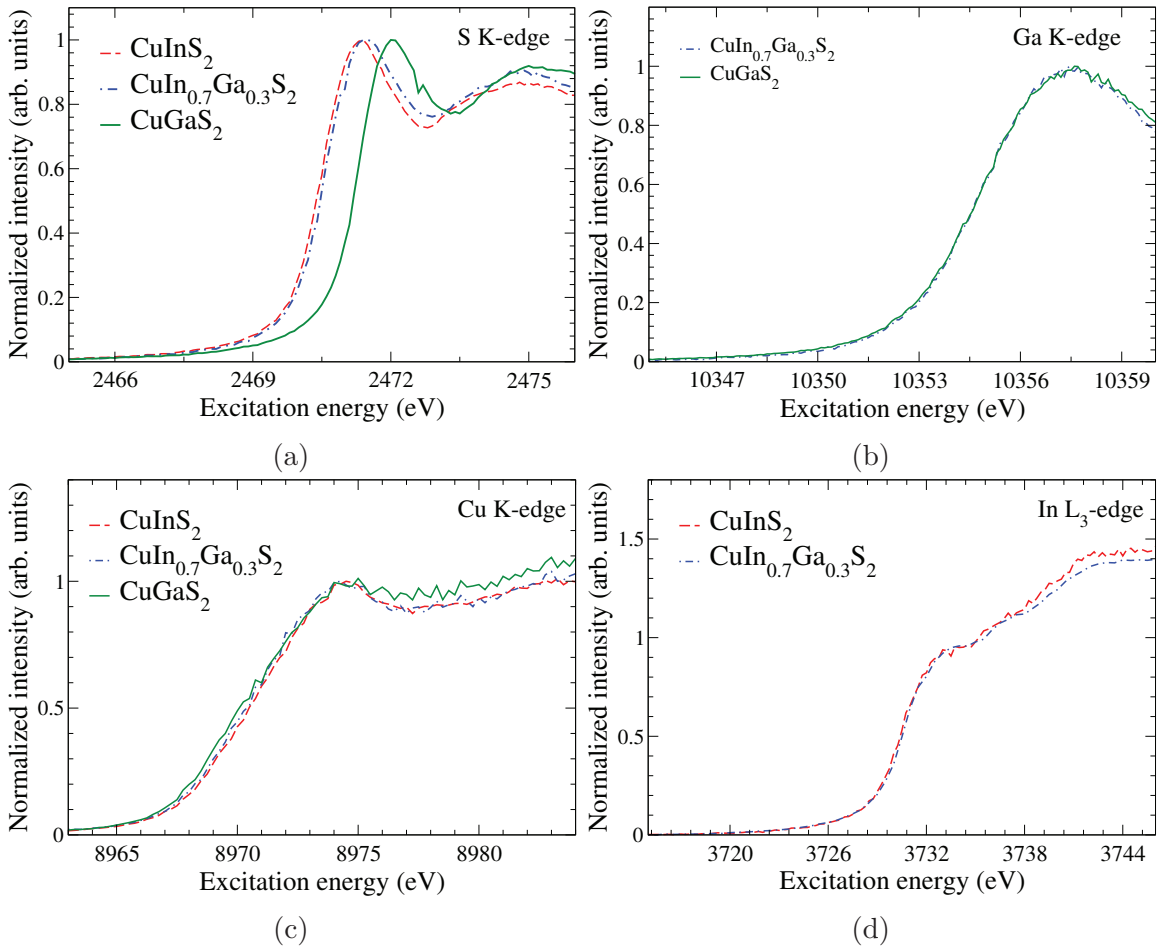


Fig. 3.8 Experimental results for samples of $CuInS_2$, $CuGaS_2$ and $CuIn_{0.7}Ga_{0.3}S_2$, from Reference [131]. a) Measured absorption K-edges of S, b) Measured absorption K-edges of Ga, c) Measured absorption K-edges of Cu, d) Measured absorption L_3-edges of In

supercells were relaxed using the all-electron projector augmented wave (PAW) method as implemented in the code VASP [37, 38]. Brillouin zone integrals were converged with a 340 eV plane-wave cutoff and a $3 \times 3 \times 3$ shifted k-point mesh. We tested different exchange-correlation potentials: Perdew–Burke–Ernzerhof (PBE)[12], PBE+U [135] with U=6 eV on Cu *d* states, and the screened hybrid Heyd-Scuseria-Ernzerhof (HSE) [81, 82]. For the ternary materials, it is already known that anion–cation bond lengths are poorly estimated when using standard semi-local functionals (such as PBE) [136–138] to relax the geometry, causing an underestimation of the band gap by more than 50% at the relaxed structure. In the mixed cation system, three cation species lead to three independent anion–cation bond lengths with two different displacement mechanisms [139].

The partial density of states (pDOS) were calculated both site- and orbital-projected using the quick projection scheme[1] implemented in VASP. In the following the pDOSs are calculated for each atom site, and for the angular quantum numbers l=0, 1 and 2 (s,p and d), summing over all the corresponding magnetic quantum numbers m. In the plots we present the sum of the pDOSs over all atom sites of the same specie.

We first checked the convergence with the cell size of the SQS. After relaxing all the SQS cells, we observed that SQS 40-atoms supercells already give bond lengths that differ by less than 0.002 Å and densities of states substantially identical to those of the 64-atoms supercells. All results shown in the following were obtained therefore with the 40-atoms SQS cells.

A further validation of our model structures comes from the comparison with Extended X-Ray Absorption Fine Structure (EXAFS) measurements of CIGS alloys performed by our collaborators. The EXAFS spectrum is caused by the interference of the electrons that are extracted out of the atoms by the incident X-rays with the scattered waves from the atoms around them. This yield the element-specific atomic-scale structure, in our case, the Cu-S, Ga-S, and In-S distances [139] (see Table 3.4).

[1]The DOS is defined as

$$D(E) = \sum_n \langle \psi_n | \psi_n \rangle \, \delta(E - E_n),$$

where, within the DFT framework, ψ_n are the KS orbitals. Inserting a complete orthonormal basis, this expression can be rewritten as

$$D(E) = \sum_n \sum_i \langle \psi_n | \xi_i \rangle \langle \xi_i | \psi_n \rangle \, \delta(E - E_n) = \sum_n \sum_i | \langle \xi_i | \psi_n \rangle |^2 \delta(E - E_n).$$

The factor $\langle \xi_i | \psi_n \rangle$ weights the terms of the DOS and defines the pDOS $D_i(E)$ as

$$D(E) = \sum_i D_i(E) = \sum_i \left[\sum_n | \langle \xi_i | \psi_n \rangle |^2 \delta(E - E_n) \right].$$

The PAW method offers a simple way of calculating those weights. The method itself needs partial waves for which atomic orbitals are used commonly which is precisely what is needed to compute the projections. Thus, within the PAW framework the partial waves ϕ_i can be used as basis ξ_i, were $i \equiv \alpha, n, l, m$, the atom site and the principal, angular and magnetic quantum numbers.

$$\langle \phi_i | \psi_n \rangle = \langle \widetilde{\phi}_i | \widetilde{\psi}_n \rangle + \langle \widetilde{\phi}_i | \hat{O} | \widetilde{\psi}_n \rangle .$$

where $\hat{O} = 1 + \sum |\widetilde{p}_i\rangle \left[\langle \phi_i | \phi_j \rangle + \langle \widetilde{\phi}_i | \widetilde{\phi}_j \rangle \right] \langle \widetilde{p}_j |$ is the overlap operator [140], $\langle \widetilde{p}_i |$ are the projectors and $\widetilde{\phi}_i$, $\widetilde{\psi}_i$ are the partial waves and pseudo KS orbitals respectively.

The second term contains $\langle \phi_i | \widetilde{p}_{i'} \rangle$ terms, which are small for i corresponding to different atom sites. i.e. $\alpha \neq \alpha'$. Thus one can approximate $\langle \phi_i | \psi_n \rangle \approx \langle \widetilde{\phi}_i | \widetilde{\psi}_n \rangle$. The final expression to compute the pDOS is

$$D_i(E) = \sum_n | \langle \widetilde{\phi}_i | \widetilde{\psi}_n \rangle |^2 \delta(E - E_n).$$

Comparing Tables 3.4 and 3.5, it can be seen that the atomic positions obtained with the HSE functional always yield bond lengths in excellent agreement with experimental data (see Table 3.5). In contrast, PBE and, to a smaller extent, PBE+U overestimates Ga-S and In-S bond lengths, while Cu-S bond lengths remain quite close to HSE and experimental values. However, all three bond lengths, Cu-S, Ga-S, and In-S, should be as close as possible to experimental values as their combination determines the anion displacement. Our findings for alloy systems are thus consistent with previous results for the pure ternaries [136, 139] and demonstrate that the use of HSE is crucial to correctly reproduce the atomic-scale structure.

Table 3.4 Experimental bond lengths of CIGS obtained from EXAFS data (Å) reported in Reference [139]

Cu/(In+Ga)	In/(In+Ga)	$d_{\mathbf{Cu-S}}$	$d_{\mathbf{Ga-S}}$	$d_{\mathbf{In-S}}$
0.99	1.00	2.325	-	2.463
0.93	0.80	2.324	2.299	2.458
0.90	0.75	2.325	2.298	2.461
0.97	0.40	2.320	2.293	2.452
1.10	0.30	2.318	2.292	2.446
1.07	0.22	2.318	2.291	2.442
0.99	0	2.312	2.286	-

Table 3.5 Calculated averaged bond lengths (Å) of the 40-atoms of $CuIn_xGa_{1-x}S_2$ supercells using different exchange correlation potentials.

	PBE			PBE+U			HSE		
x	$d_{\mathbf{Cu-S}}$	$d_{\mathbf{Ga-S}}$	$d_{\mathbf{In-S}}$	$d_{\mathbf{Cu-S}}$	$d_{\mathbf{Ga-S}}$	$d_{\mathbf{In-S}}$	$d_{\mathbf{Cu-S}}$	$d_{\mathbf{Ga-S}}$	$d_{\mathbf{In-S}}$
1.00	2.33	-	2.52	2.33	-	2.50	2.34	-	2.48
0.70	2.33	2.34	2.51	2.33	2.33	2.50	2.34	2.31	2.47
0.50	2.32	2.34	2.50	2.33	2.33	2.49	2.34	2.30	2.46
0.30	2.32	2.33	2.49	2.33	2.32	2.48	2.33	2.29	2.46
0.00	2.31	2.32	-	2.32	2.31	-	2.32	2.29	-

Besides, the DOSs of the relaxed structures show that the HSE functional yields a band gap very close to the experimental value [136, 141] (see Table 3.6). Hence, to have all three bond lengths, Cu-S, Ga-S, and In-S, as close as possible to experimental values (as their combination determines the anion displacement) and a better electronic structure, HSE was used in this work.

In order to evaluate band offsets, it was necessary to align the band edges of compounds of different chemical compositions. For this, we used the method described in the work of Schleife *et al.* [142] and aligned the branch-point energies, defined as the energy at which the defect states induced in the gap change their character from predominantly acceptor-like to donor-like [143, 144]. Reaching from one material into another at the interface such states transfer a net charge, the sign of which depends on the position of the Fermi level relative to the branch point energy. This charge transfer leads to an intrinsic interface dipole that tends to line up the energy bands in a way that the dipole itself vanishes [144]. Therefore, branch point energies are relevant reference levels for the band alignment, and their use allows us to avoid the direct calculations of interfaces.

The branch point energy E_{BP} is calculated as a Brilluoin Zone average of the KS eigenvalues of the lowest N_{CB} conduction bands and the highest $N_{\mathrm{VB}} = 2N_{\mathrm{CB}}$ valence [142, 145].

$$E_{\mathrm{BP}} = \frac{1}{2N_{\boldsymbol{k}}} \sum_{\boldsymbol{k}} \left[\frac{1}{N_{\mathrm{CB}}} \sum_{i}^{N_{\mathrm{CB}}} \varepsilon_{c_i}(\boldsymbol{k}) + \frac{1}{N_{\mathrm{VB}}} \sum_{j}^{N_{\mathrm{VB}}} \varepsilon_{v_j}(\boldsymbol{k}) \right].$$

This method requires to determine the number of bands for the calculation of E_{BP}. Following the scaling of the number of bands with the number of electrons described in Reference [142], we obtain for $CuGaS_2$ and $CuInS_2$, with 32 electrons on the unit cell, 8 valence bands and 4 conduction bands and for $CuIn_xGa_{1-x}S_2$ supercells with 160 electrons, 40 valence bands and 20 conduction bands.

With the valence band maximum E_v=0 as energy zero, the conduction band minimum E_c takes the value of the band gap E_g. From E_{BP} and E_g the band offsets are calculated as

$$\Delta E_c = [E_g(X_1) - E_{\mathrm{BP}}(X_1)] - [E_g(X_2) - E_{\mathrm{BP}}(X_2)],$$

$$\Delta E_v = E_{\mathrm{BP}}(X_1) - E_{\mathrm{BP}}(X_2),$$

where X_1 and X_2 are the two compounds forming the interface.

Figure 3.9 shows the calculated pDOS of the S, Ga and Cu p-states and the In s+d-states, for the conduction bands of CIGS with varying composition. A Gaussian broadening of 0.3 eV is applied to all the calculated pDOSs. According to the dipole selection rules, which states that the change in angular momentum quantum number should be $\Delta l = \pm 1$ between the initial and final states, electrons from the K-shell

Table 3.6 Calculated and measured valence band offsets ΔVBM and band gaps E_g of $CuIn_xGa_{1-x}S_2$ alloys (eV). Experimental values are from Ref. 3

	ΔVBM		E_g	
x	HSE	Exp.	HSE	Exp.
1.0	0.14	0.3	1.32	1.5
0.7	0.11	0.3	1.55	1.6
0.5	0.10	-	1.69	-
0.3	0.06	-	1.90	-
0.0	0.00	0.0	2.20	2.4

(s-orbitals) are excited to unoccupied conduction bands with p character. Furthermore, transitions only take place if the final and initial states overlap in space [146]. Since the core level states, particularly of the K-shell, are highly localized at the absorbing atom, the electrons will be excited to those unoccupied states whose amplitudes dominate at the absorber site: the states of the absorbing atom itself. Thus, the pDOSs of the S, Ga and Cu p-states and the In s+d-states of Figure 3.9 correspond to the unoccupied states sampled at the S, Ga and Cu K-edges and In L_3-edge (Figure 3.8) respectively.

The qualitative agreement between calculated and measured pDOS is excellent as can be seen from comparing Figure 3.8 and Figure 3.9. In particular, a significant shift of the edge position with alloy composition is observed for S in both cases, whereas the Ga, In and Cu edges position are less dependent on composition for both experiment and calculation. The theoretical findings thus confirm the experimental observations demonstrating that this unusual behavior of absorption edge positions is indeed a real effect of mixed chalcopyrite alloys.

The fact that the Ga, In, and Cu absorption edge positions remain nearly unaltered with alloy composition despite a significant change of the band gap closely resembles the fact that the element-specific Ga-S, In-S, and Cu-S bond lengths are nearly constant over the whole compositional range despite a significant change of the lattice constants. In contrast, both the S-edge position and the average group-IIIA-S bond length vary with changing In/Ga ratio. This strongly suggests that the absorption edge position, i.e., the energy of the element-specific unoccupied local states, is determined by the local atomic arrangements rather than the overall crystallographic structure.

In the chalcopyrite each Ga, In, or Cu cation is bonded to four S anions. If all atoms occupied ideal lattice sites, the Ga-S and In-S bond lengths would be identical and would change with alloy composition according to the change in lattice constants.

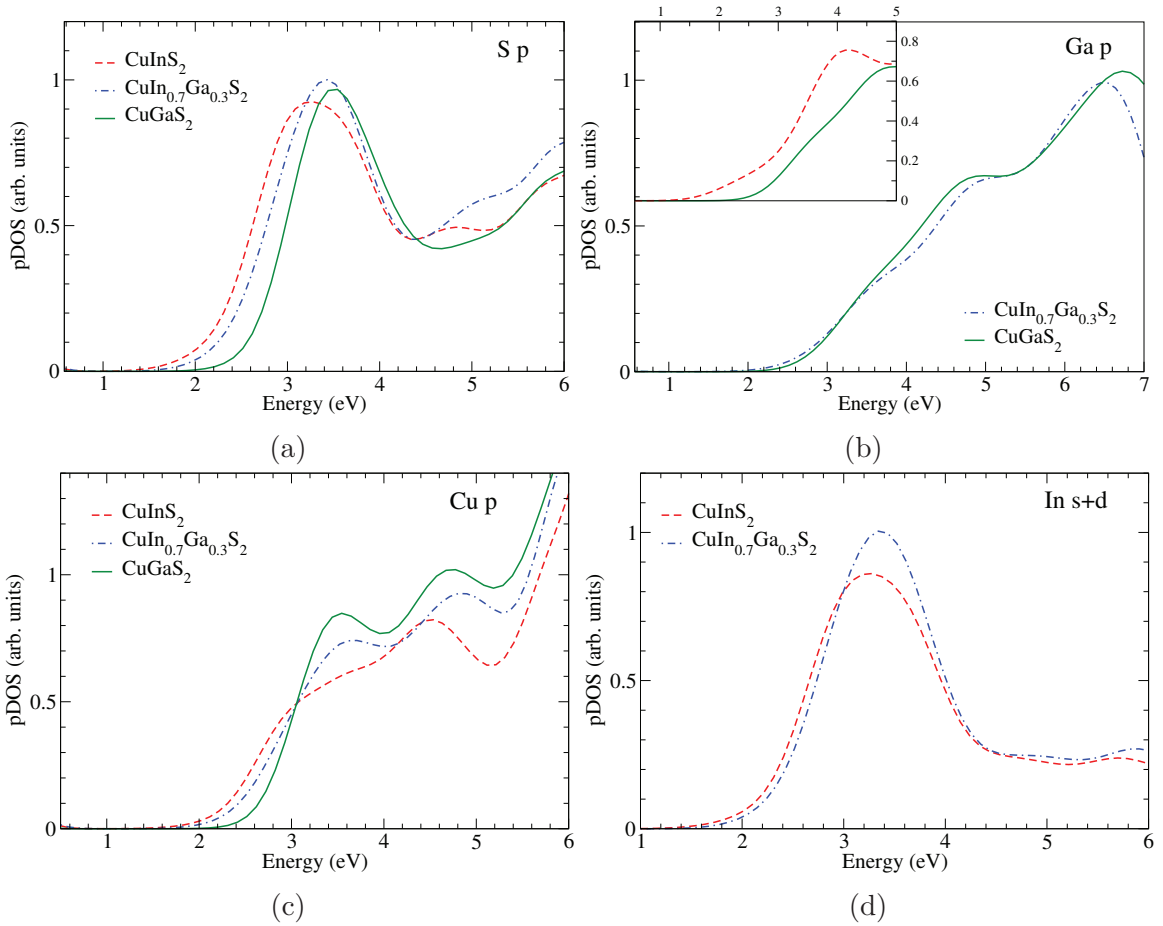


Fig. 3.9 Calculated pDOS for the conduction bands to be compared with Figure 3.8. The valence band maximum is set to zero and the energies scales of different compounds are aligned via the branch point energy. a) *p*-states of S, b) *p*-states of Ga. Inset: pDOS of Ga *p*-states in the conduction band for CGS with either CGS lattice constants (green solid line) or with average lattice constants between CIS and CGS (red dashed line). c) *p*-states of of Cu, d) *s*-states plus *d*-states of In

Note that the Cu-S bond length is nearly the same for CGS and CIS and hence no significant structural change is expected for the alloy. The local structural environment surrounding the Ga or In atoms would thus be expanded or compressed leading to a change in band gap according to the well-known dependence on pressure or temperature. Consequently, the absorption edge position would shift with varying alloy composition. This effect is illustrated in the inset of Figure 3.9b where the calculated pDOS of Ga *p*-states in the conduction band is plotted for CGS with either the real CGS lattice constants or with those applying to the x=0.5 alloy. The edge position clearly shifts as the Ga-S bond length is stretched from the ternary length to the distance of the alloy lattice sites.

Obviously, this does not correspond to the behavior of the real CIGS alloys as shown in Figure 3.9 with the reason being that the local atomic arrangements strongly deviate from the average crystallographic structure. EXAFS measurements of both CIGS and CIGSe have shown that the element-specific bond lengths are nearly constant over the whole compositional range despite the change in lattice constants [132, 139].This behavior closely resembles the findings for other mixed semiconductor systems as first reported for (In,Ga)As [147] and later confirmed for many other IIIA-VA and IIA-VIA ternary compounds[148, 149]. It originates from the fact that bond bending is energetically favored over bond stretching such that the lattice mismatch is accommodated in the mixed system mostly by a change of the bond angles and only to a small extent by a change of the bond lengths [148, 150]. In mixed cation systems, this is achieved by a displacement of the anion from its ideal lattice site [151], which has been shown to influence the band gap of the material for both IIIA-VA alloys [152] and mixed chalcopyrites [132, 139]. This demonstrates again the strong similarity between these tetrahedrally coordinated mixed semiconductors and highlights the correlation between atomic-scale structure and electronic properties. However, due to the increased complexity of the chalcopyrites compared to the IIIA-VA or IIA-VIA compounds, two different displacement mechanisms must be distinguished for CIGS and CIGSe as discussed in detail in Reference [132, 139].

For understanding the behavior of the absorption edge positions discussed here, the important feature of the atomic scale structure is the fact that the element-specific Ga-S, In-S, and Cu-S bond lengths remain close to the ternary values over the whole compositional range. As a consequence of this bond length preservation, the local structural environment surrounding the Ga, In, or Cu atoms changes very little and the projected pDOS still resembles that of the pure ternary compounds. The corresponding absorption edge position is thus independent of the alloy composition and remains fixed at the ternary energy position.

In contrast, the S anions are bonded to two Cu and two group-IIIA atoms which can be either Ga or In. Consequently, the average first nearest neighbor environment of S changes with changing alloy composition even if the element-specific bond lengths remain constant. The S pDOS in the alloy thus represents a weighted average of the S pDOS in CGS and CIS corresponding to an absorption edge position that shifts with changing In/Ga ratio as seen in Figure 3.8. The change of the band gap, i.e., the shift of the conduction band minimum, is thus caused by a changing spatial average over the element-specific local states rather than by a change in energy of these states themselves. As a consequence, a determination of the conduction band minimum and

thus the band gap from X-ray absorption spectroscopy studies is not straightforward in semiconductor alloys and care has to be taken when evaluating such data.

A very similar behavior can be found in Yamazoe *et al.* where Cu-In-Se compounds were studied with different Cu/In ratios varying from 1 ($CuInSe_2$) to 0.2 ($CuIn_5Se_8$) [153]. While the Cu and In K-edge NEXAFS are independent of the sample stoichiometry, the Se K-edge NEXAFS exhibits significant changes as the Cu/In ratio decreases. Despite the different crystal structures of the various compounds, the Cu and In cations are bonded to four Se anions the distance to which varies only little with stoichiometry. The local structural environment of Cu and In, and hence the pDOS, therefore, remains mostly unchanged. In contrast, the average surrounding of the Se anions changes strongly with changing Cu/In ratio as clearly evidenced by the changing NEXAFS spectra. In view of these similarities between CIGS and Cu-In-Se compounds and given the remarkable resemblance of the atomic-scale structure of CIGS and CIGSe [132, 139], we strongly believe that the findings presented in this work are general features of tetrahedrally coordinated semiconductors with different cation species and varying composition or stoichiometry. The similarity between our study and those on $La_{1-x}S_xCoO_3$ [154–156] further suggests that this behavior is even more general and can be found whenever an alloy contains elements for which the local first nearest neighbor environment does or does not change with changing alloy composition. For both CIGS and $La_{1-x}S_xCoO_3$, the nature and distance of the anions surrounding the cations do not change and the cation absorption edge position does not shift. In contrast, the local environment surrounding the anions does change with alloy composition and consequently the anion absorption edge positions do shift in both cases.

In conclusion, DFT calculations together with the results of NEXAFS measurements show that the absorption edge position for Ga, In, and Cu is independent of the composition, whereas the S absorption edge shifts with changing In/Ga ratio in accordance with the change in bandgap. This behavior originates from the state selectivity of the absorption process in which the core level electron is excited predominantly into unoccupied states of the absorbing atom itself. These element-specific local states are determined by the atomic-scale structural environment, which is nearly independent of composition for the Ga, In, and Cu cations but varies significantly with changing In/Ga ratio for the S anions. The observed change in the band gap with changing alloy composition thus results from a changing spatial average of the nearly unchanged element-specific local states. This clearly demonstrates the strong influence of local structural parameters on the electronic properties of the material. For those to

be predicted correctly, the DFT calculation therefore has to reproduce not only the crystallographic structure correctly but also the atomic-scale structural parameters such as element-specific bond lengths and anion displacement which necessitates the use of the HSE exchange-correlation potential. While the study was performed for the specific case of CIGS, we believe that our findings are also applicable other compound semiconductors with different cation species and varying composition or stoichiometry.

Chapter 4

Optimized Exchange and Correlation Functional for the Calculation of Energies of Formation

In this chapter we describe how we develop a semi-empirical exchange-correlation functional for Density Functional Theory, tailored to calculate energies of formation of solids. This functional has the form of a Perdew-Burke-Ernzerhof functional, but with three parameters, covering the exchange and correlation parts, fitted to reproduce experimental energies of formation for a representative set of binary compounds. The quality of the obtained functional was then assessed for a control set. Our functional manages to reduce the error of the Perdew-Burke-Ernzerhof generalized gradient approximation by roughly a factor of two. Furthermore, this is achieved without affecting the quality of the geometry.

4.1 Introduction

The accuracy of many thermochemical experiments, the so-called chemical accuracy, is commonly taken as approximately 40 meV/formula-unit (1kcal/mol). This is the desirable accuracy of theoretical methods as well. To achieve such a level of accuracy one needs to pay the price of very high computational cost. In fact, calculating energies of formation at the chemical accuracy level is nowadays possible only by using very expensive (and therefore limited to very small systems) first-principles methods such as Configuration-Interaction or Quantum Monte Carlo.

Density Functional Theory (DFT) could give in principle exact ground-state energies (in the Born-Oppenheimer framework), but in practice it does not, as the exact form of the exchange-correlation functional is unknown. In particular, the common semi-local approximations for the exchange-correlation energy term usually suffer from the self-interaction problem, i.e. the interaction of an electron with itself. The self-interaction error is believed to be the cause of many failures of these functional. DFT is widely used in structure prediction, i.e. to compute energies of formation, as it is the only approach that allows to handle the large number of calculations required by large scale calculations

[19, 157, 158] obtaining a reasonable compromise between accuracy and computational effort. DFT energies of formation (in the local density approximation or generalized gradient approximation) often benefit from error cancellations when subtracting the total energy of the elemental phases to the total the energy of the compounds. This occurs in a larger extent when considering energy differences between similar systems, for example, metallic alloys and their (metal) components [159]. Conversely, DFT is more likely to fail to reproduce the energies of formation of semiconductors and insulators composed by metals and molecular species [159].

There are some other sources of errors involved when it comes to predict a thermochemical quantity, even at zero temperature. Namely, originating from the zero-point motion of the phonons. However, and as in the vast majority of cases, the error due to phonons is considerably smaller than the error due to the exchange-correlation, therefore we will ignore it in the following. Extrapolated experimental values to 0 K, show that the errors introduced by finite temperature are typically smaller than 0.03 eV/atom [159]. This makes possible the comparison to the available experimental enthalpies of formation, which usually correspond to standard conditions: temperature of 298 K and pressure of 101.3 kPa. On the other hand, one should not forget that the contribution of the zero-point motion of diatomic molecules O_2 (0.049 eV/atom), N_2 (0.073 eV/atom) and H_2 (0.135 eV/atom) is significant [160].

Several approaches have been suggested to improve the prediction of the energies of formation. Most of them include a Hubbard U parameter which allows to better describe the *d*-electrons and hence to obtain better total energies [157, 159, 161]. However, there are no unique rules to choose its values. This is in general a "thermodynamical" U, in the sense that it is not intended to predict band-gaps, but instead total energies leading to accurate energies of formation.

In turn, there are several ways to correct the energies of formation obtained in the DFT+U framework. Lei Wang *et al.* [157] computed the formation of transition metal oxides with the PBE and PBE+U methods, and found a constant shift of 1.36 eV per O_2 molecule between the experimental and theoretical energies of formation. This comes in part from the inaccuracy of the PBE exchange-correlation functional in describing the binding energy of the O_2 molecule which is about 1.58 eV per O_2 molecule lower than the experimental value. The remaining contribution to the shift is attributed to the PBE error associated with adding electrons to the O *p*-orbital when ions O^{2-} are formed from O_2. They also provide the values of U for several transition metals, that give the best oxidation energies.

Anubhav *et al.* [161] presented a way to mix PBE and PBE+U total energies to obtain more accurate energies of formation. Their methodology is based on the idea of decomposing the reaction under study as a set of subreactions that either are well described with PBE alone, well described with PBE+U alone, or for which there is experimental data available. The subreactions will introduce intermediate phases, whose energies are computed with either PBE or PBE+U. For a set of 49 ternary oxide, this approach was able to reduce the mean absolute relative error of calculated formation enthalpies from approximately 7.7–21% in PBE + U to less than 2%. They also successfully predict that all the known phases of the Fe-P-O system are stable in contradiction with PBE and PBE+U calculation.

In a completely different approach, Stevanovic *et al.* [159] used a set of 252 binary compounds to determine corrections to the energy of the elements in order to maximize the error cancellation leading to accurate values for the energy of formation. They wrote the energy of each elemental phase as $E^{\mathrm{PBE+U}}(A)+\mu_A$ and use μ_A as a parameter to fit the PBE+U-computed formation energy to the experimental value. By using these fitted elemental-phase reference energies (FERE), the energies of formation of a set of 252 binaries are reproduced with the mean absolute error of 0.054 eV/atom instead of 0.250 eV/atom resulting from pure PBE calculations. Using FERE in a set of 55 ternary compounds that were not part of the fitting set, the energies of formation are reproduced with a mean absolute error of = 0.048 eV/atom. Their choice for the values of the Hubbard U was to set it to 3 eV for all transition metals except Ag and Cu, which was set to 5 eV. The strategy of fixing U allows to develop a more general scheme that can be applied to different families of compounds, not only to oxides or chalcogenides separately, for instance. The μ_A corrections, added to the elemental-phase energies, automatically take care of the contribution of the zero-point motion of the diatomic molecules mentioned above.

This is probably the best method available for the calculation of energies of formation, but unfortunately it has a few shortcomings when used to obtain the distance to the convex hull of stability. To understand these shortcomings, we should remember that to calculate the distance to the convex hull we need to consider all possible decomposition channels. It turns out that depending on the channel, FERE can give very different results: (i) The decomposition is to elementary phases. Then the distance to the hull is simply related to the energy of formation and FERE yields excellent results. (ii) The solid decomposes into one (or more) elementary phases and one (or more) binary or higher compounds. In this case FERE just corrects the elementary phases, and the error can be substantially higher than for the normal PBE+U. This

can be easily seen in, e.g., binary phases diagrams as FERE can (incorrectly) stabilize (or destabilize) certain phases. (iii) Otherwise the decomposition is into binaries or higher phases, and FERE simply yields the same error as the underlying PBE+U theory.

In this chapter we present a new approach to improve the accuracy of DFT thermochemistry: a generalized gradient approximation exchange-correlation functional optimized to give accurate values of energies of formation. To define the new functional we tuned the values for the parameters μ, κ, β and γ of the PBE [12] to fit experimental formation energies. We also verified that with the optimized functional, the lattice parameters are still determined with the same accuracy of PBE, which is an essential requirement to use it for structure prediction.

4.2 Generalized Gradient Approximation for the Exchange and Correlation energy

The challenge in the Kohn-Sham DFT formalism is to describe the exchange-correlation term of the total energy. Often one divides this term into separate exchange and correlation contributions [4, 5]

$$E_{\mathrm{xc}} = E_{\mathrm{x}} + E_{\mathrm{c}}. \tag{4.1}$$

The generalized gradient approximation (GGA) for exchange is usually expressed as [4, 5, 12]

$$E_{\mathrm{x}}^{\mathrm{GGA}}[n] = \int d\boldsymbol{r} n \epsilon_{\mathrm{x}}^{\mathrm{unif}}[n] F_{\mathrm{x}}(s), \tag{4.2}$$

(although non-separable forms have also been proposed) where $s = |\nabla n|/2k_F n$ is a dimensionless density gradient, F_{x} is called enhancement factor and $\epsilon_{\mathrm{x}}^{\mathrm{unif}}(n) = -3k_F/4\pi$ is the exact uniform-electron-gas exchange energy and $k_F = [3\pi^2 n(\boldsymbol{r})]^{1/3}$ is the local Fermi wavevector.

We will write the GGA for the correlation in the form [4, 5, 12]

$$E_{\mathrm{c}}^{\mathrm{GGA}}(n_\uparrow, n_\downarrow) = \int d\boldsymbol{r} n [\epsilon_{\mathrm{c}}^{\mathrm{unif}}(r_s, \zeta) + H(r_s, \zeta, t)], \tag{4.3}$$

where H is the gradient contribution, r_s is the Seitz radius, defined as $n = 3/4\pi r_s^3 = k_F^3/3\pi^2$, $\zeta = [n_\uparrow - n_\downarrow]/n$ is the relative spin polarization, and $t = |\nabla n|/2\phi k_s n$ is a dimensionless density gradient, with $\phi(\zeta) = [(1+\zeta)^{2/3} + (1-\zeta)^{2/3}]/2$, a spin-scaling factor, and $k_s = \sqrt{4k_F/\pi}$, the Tomas-Fermi wavenumber.

The analytic expressions for the correlation energy of the uniform electron gas ϵ_c are known only in the extreme limits when $r_s \to 0$ and $r_s \to \infty$. An expression satisfying both limits, proposed by Perdew and Wang [4, 6, 8], is

$$\epsilon_c(n) = -2c_0(1+\alpha_1 r_s)\ln\left[1+\frac{1}{2c_0(\beta_1 r_s^{1/2}+\beta_2 r_s+\beta_3 r_s^{3/2}+\beta_1 r_s^2)}\right], \tag{4.4}$$

where $c_0 = 0.031091$, $\alpha_1 = 0.21370$, $\beta_3 = 1.6382$, $\beta_4 = 0.49294$, $\beta_1 = 1/(2c_0)\exp(-c_1/2c_0)$, $\beta_2 = 2c_0\beta_1^2$, where $c_1 = 0.0466644$. These coefficients were found by fitting to Quantum Monte Carlo uniform electron gas correlation energies for r_s=2, 5, 10, 20 and 100 [7].

An enhancement factor for correlation can be defined as [5]

$$F_c(r_s,\zeta,t) = 1 + \frac{H(r_s,\zeta,t)}{\epsilon_c^{\text{unif}}(r_s,\zeta)}, \tag{4.5}$$

allowing us to write the correlation energy as

$$E_c^{\text{GGA}} = \int d^3\boldsymbol{r} n \epsilon_c^{\text{unif}} F_c(r_s,\zeta,t). \tag{4.6}$$

Equivalently, an enhancement factor for the exchange and correlation together can be defined as

$$F_{xc}(r_s,\zeta,s) = F_x(s) + \frac{\epsilon_c^{\text{unif}}(r_s,\zeta)}{\epsilon_x^{\text{unif}}(r_s,\zeta)} F_c(r_s,\zeta,s) \tag{4.7}$$

and the exchange-correlation energy can be written as

$$E_{xc}^{\text{GGA}} = \int d^3\boldsymbol{r} n \epsilon_x^{\text{unif}} F_{xc}(r_s,\zeta,t). \tag{4.8}$$

4.3 The PBE Exchange-Correlation Functionals Family

The Perdew-Burke-Ernzerhof (PBE) [12] is the most used exchange-correlation functional in DFT calculations in condensed matter and it has a simple expression very easy to derive and implement. The PBE approximation to correlation is constructed from three conditions [12]:

a) at the slowly varying limit $t \to 0$, the gradient contribution to the correlation is given by the second-order expansion

$$H \to \beta\phi^3 t^2, \tag{4.9}$$

where $\beta \simeq 0.066726$.

b) at the rapidly varying limit $t \to \infty$ the correlation vanishes

$$H \to -\epsilon_{\rm c}^{\rm unif}, \tag{4.10}$$

c) due to the uniform scaling (1.58) to the high density limit the correlation energy must scale to a constant, thus H must cancel the logarithmic singularity of $\epsilon_{\rm c}^{\rm unif}$.

The selected function to satisfy a) b) and c) is the first term of the Perdew-Wang-91 (PW91) correlation functional [162]

$$H = \gamma\phi^3 \ln\left[1 + \frac{\beta}{\gamma} t^2 \frac{1 + At^2}{1 + At^2 + A^2t^4}\right], \tag{4.11}$$

where

$$A = \frac{\beta}{\gamma} \exp\left(-\frac{\epsilon_{\rm c}^{\rm unif}}{\gamma\phi^3} - 1\right)^{-1}, \tag{4.12}$$

The GGA for the exchange is constructed using four further conditions [12]

d) in the uniform density scaling (1.58), $E_{\rm x}$ must have the form (4.2) to recover the correct uniform electron gas limit with $F_{\rm x}(s = 0) = 1$

e) the exact exchange energy obeys the spin-scaling relation

$$E_{\rm x}[n_\uparrow, n_\downarrow] = \frac{E_{\rm x}[2n_\uparrow] + E_{\rm x}[2n_\downarrow]}{2}\ , \tag{4.13}$$

f) it is desired to recover the local spin density linear response when $s \to 0$, hence the effective gradient for exchange

$$F_{\rm x}(s) \to 1 + \mu s^2\ , \tag{4.14}$$

has to cancel that for correlation, leading to the condition $\mu = \beta\pi^2/3 \simeq 0.21951$.

g) the Lieb-Oxford bound [13]

$$E_{\rm x}[n_\uparrow, n_\downarrow] \geq E_{\rm xc}[n_\uparrow, n_\downarrow] \geq -1.679 \int d^3\boldsymbol{r} n^{4/3}\ , \tag{4.15}$$

will be satisfied if $F_{\rm x}(\zeta = 1, s)$ grows gradually with s to a maximum value less than or equal to 2.273.

A simple function satisfying d), e), f) and g) was proposed by Becke [163] but with empirical coefficients

$$F_{\rm x}^{\rm PBE}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa}\ . \tag{4.16}$$

Table 4.1 The parameters μ, κ, β and γ for some functionals of the PBE family.

	PBE	PBEsol	xPBE	revPBE	B86
μ	0.21951	10/81	0.23214	0.21951	0.23511
κ	0.804	0.804	0.91954	1.245	0.9672
β	0.066725	0.046	0.089809	0.066725	-
γ	0.031090	0.031090	0.011279	0.031090	-

Although the original derivation of the PBE used solid theoretical arguments to fix the four parameters μ, κ, β and γ, there seems to still be a considerable amount of other (reasonably well) justified choices. In different works, these parameters have been modified to improve the performance of the original PBE functional to predict magnitudes as atomic data, heats of formation of molecules and lattice parameters. That gives rise to a family of functionals based on PBE among which are PBEsol [80], xPBE [5], revPBE [164], PBE-TCA [165, 166], PBEA [167] ,vdW-PBE [168], RPBA [169], PBEint [170].

The functional revPBE, whose name stands for revised PBE, appears in a comment to the paper where the PBE functional is presented. Zhang *et al.* [164] argue that the derivation of PBE leaves flexibility to choose κ and they obtain a new value of this parameter by fitting exchange-only total energies of atoms to exact exchange-only results from the optimized exchange potential method, finding that it can improve significantly over the original PBE. The revPBE improves PBE atomic total energies by a factor of 10 [164]. In the PBEsol functional, the gradient expansion is restored to better describe real solids, where the density is often almost slowly varying over space. The β parameter is chosen to best fit the results with the Tao-Perdew-Staroverov-Scuseria (TPSS) functional [171] for large neutral jellium clusters. To construct the extended PBE functional xPBE, the four parameters are fitted to the Hartree-Fock limit energies, to the exact atomic energies for atoms from H to Ar, the binding energies, and bond distance of Ne_2. xPBE outperforms PBE for thermochemistry of molecules, and can be competitive or better than PBE for the prediction of geometric parameters, ionization potentials, electron affinities and for the description of Van der Waals and hydrogen bond interactions [5].

4.4 OPTIMIZED SEMI-LOCAL EXCHANGE-CORRELATION FUNCTIONAL FOR THE CALCULATION OF ENERGIES OF FORMATION

To optimize our exchange-correlation functional, we minimized the mean absolute deviation (MAD) of the PBE energies of formation of a set of 92 binary compounds with respect to the experimental values, obtaining a fit for the four parameters μ, κ, β and γ on (4.11) and (4.16).

The optimization set was selected to assure its chemical diversity. It is a subset of the 252 binary compounds used on Reference [159], for which experimental energies of formation are known. It comprises 47 different elements (see Table C.5) and it contains 44 compounds of transition metals, 38 compounds of alkali and alkali-earth metals and 12 compounds of the groups IIIA and IVA. They are listed in Table C.1. Structures were obtained from the Material Project database [127]. We included as many distinct chemical elements as possible, and that were representatives of all groups of the periodic table. We could take a step beyond, and also require representative structures for all possible oxidation states for all elements, for example. This would, certainly, cover even better the chemical diversity. However, this would lead to a too large training set and often too large unit cells for our current computational ability.

The ground state structure of the elemental solids at 0 K, were obtained from Reference [172] or from the Material Project Database [127]. The structures were relaxed using PBE functional and the ones with lowest energies were used for the computation of the energies of formation.

The minimization was performed with the Downhill Simplex method (DSM) by Nelder and Mead [173], being the parameters μ, κ, β and γ, the variables, and the MAD, the objective function. The DSM requires only function evaluations, not derivatives. Several minimization were performed starting from different initial parameters in order to check the convergence of the method. This method uses the concept of simplex: the geometrical figure consisting of N + 1 vertices and all their interconnecting line segments, polygonal faces, etc, where N is the number of variables of the function. The algorithm starts with N + 1 points, defining an initial simplex. Then it takes a series of steps in its way downhill, most of them just moving the point of the simplex where the function is largest through the opposite face to a lower value (called reflexion). The simplex can also be contracted to enter in narrow zones.

All calculations were done within the projector augmented wave (PAW) method as implemented in the VASP code [37, 38]. The plane wave cut-off was set to a value 30% higher than the ones suggested for the PAW potentials. The numerical convergence to less than 2 meV/atom was ensured by a high energy cutoff and dense k-point meshes.

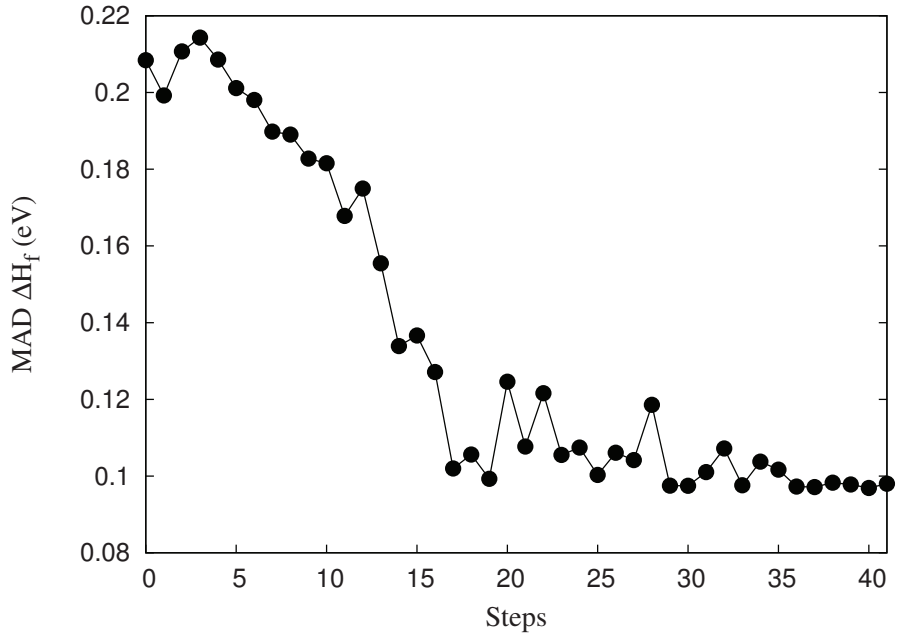


Fig. 4.1 DSM minimization of the MAD starting from the PBE values of the parameters.

For each functional the structures were relaxed until the forces over the atoms were less than 0.005 eV/Å and the stresses over the cell were lower than 0.1 GPa. All the total energies used to compute energies of formation correspond to relaxed structures.

The calculations involving transition metals were performed spin polarized, with a ferromagnetic alignment as initial spin configuration, except for NiO, CoO, CuO and FeO that where set to start from an antiferromagnetic alignment. Note that Stevanovic *et al.* [159] reported that they found that the energy differences associated with different magnetic configurations are typically of the order of 0.01-0.02 eV/atom and do not contribute appreciably to relative large *ab initio* errors.

Figure 4.1 shows a typical minimization of the MAD. The step zero corresponds to the PBE parameters for which the MAD i equal to 0.208 eV/atom. The following four steps determine the other points corresponding to the first simplex, and then, in step 5 the minimization starts.

The minimum found for the MAD is 0.097 eV/atom for parameters $\mu \simeq 0.346$, $\kappa \simeq 0.437$, $\beta \simeq 0.043$ and $\gamma \simeq 0.034$ (see Figure 4.2). These values correspond approximately to 1.5 times μ_{PBE}, 0.6 times κ_{PBE}, 0.66 times β_{PBE}, (very close to the β_{PBEsol}) and almost the same γ_{PBE}. Repeating the calculations with the converged parameters but fixing γ_{PBE}, do not change significantly the MAD (0.099 eV/atom). Therefore, we decided to leave γ unchanged.

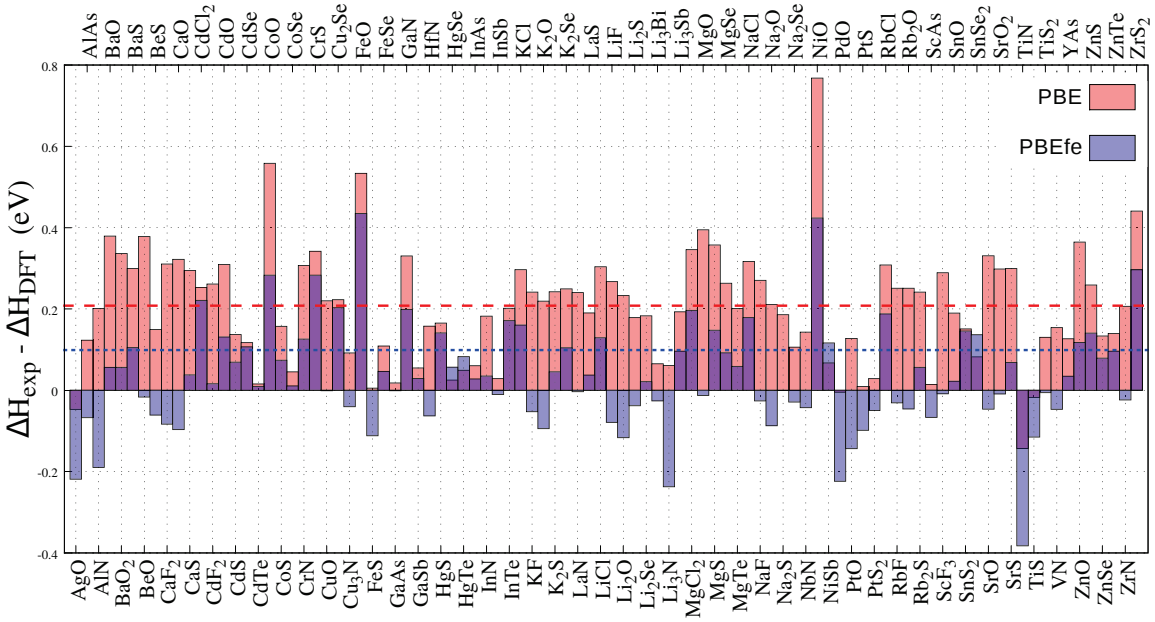


Fig. 4.2 Errors of the calculated energies of formation with the PBE and the PBEfe functionals. The orange long-dashed line represents the MAD of the PBE and the blue short-dotted line represent the MAD of PBEfe. Transparent colors have been used to show the orange bars when they are shorter than the corresponding blue one.

The individual errors for the different materials can be seen in Figure 4.2 and Table C.3. PBE formation energies are affected by systematic errors, leading to an overestimation of the experimental values for most of the compounds, i.e. values less negatives than the experimental ones. This was already pointed out by Lany [174]. The energies are very concentrated with a few outliers that correspond to highly correlated Mott insulators as NiO and CoO, which are incorrectly described by semi-local approximations. This is different if our optimized functional (which we label PBEfe standing for PBE for formation energies) is used instead: for 38 compounds, out of 92 on the list, the calculated energies of formation are underestimated, while the rest is overestimated. Table 4.2 summarizes these results. We also presented for the sake of comparison the MAD of the energy of formation computed from the total energies obtained from of the Materials Project Database, which use the Hubbard U values and the corrections to the O_2 binding energy from Reference [157].

Figure 4.3 shows a comparison between the calculated and experimental energies of formation of the 92 compounds with four different functionals, PBE, PBEsol, xPBE, revPBE, LDA and PBEfe. In addition to the overestimation, it can be seen that PBE

Table 4.2 Mean absolute deviation, and maximum absolute error (MAE) for our optimization set with the different functionals mentioned here and the results obtained in the Material Project (MP) Database.

	MAD (eV)	MAE (eV)
PBE	0.208	0.768
PBEsol	0.193	0.664
xPBE	0.225	0.631
revPBE	0.262	0.701
LDA	0.152	0.672
PBEfe	0.099	0.435
PBEfe+U	0.078	0.336
MP	0.130	0.394

errors increase with the absolute value of the total energies. The functionals xPBE and revPBE gives the largest errors increasing the MAD up a 21% with respect to the PBE. The energies obtained with the PBEsol functional are slightly better than the ones that PBE yields. LDA energies are considerably better, and the values do not have such a large systematic error as PBE or PBEsol. This was expected as it was also pointed out by Lany [174]. Finally, PBEfe results show a much smaller dispersion.

The PBEfe error of the subset of the 92 systems formed of transition-metals compounds is higher (0.120 eV/atom) than the remaining subset of the *s*- and *p*-compounds (0.077 eV/atom). To address that, we performed calculations applying a Hubbard U term to the *d*-orbitals of the transition metal compounds. To obtain the values of U we used the following procedure: (i) For each compound we found the value of U for which PBEfe+U gave the lowest error. (ii) We averaged the best values of U for all compounds containing the element. This is not intended to find universal U values, but just to compensate the difference of errors between the mentioned subsets. Table 4.3 shows the values of U, for each transition metal. The PBEfe+U calculation gives a MAD of 0.078 eV/atom. Figure 4.3 shows how the calculated energies of formation are obtained closer to experiment when applying PBEfe+U.

It is important to remark that the search of for more accurate exchange-correlation functionals will never lead to a MAD of zero. Even if the exact functional is known, the energies of formation are still going to be predicted with a considerable error. In fact, calculations in the Born-Oppenheimer framework neglects the energetic contribution of the atomic motion, which is of the order of tens of meV/atom (including zero point motions) as discussed before, that is not much smaller than the value of the MAD

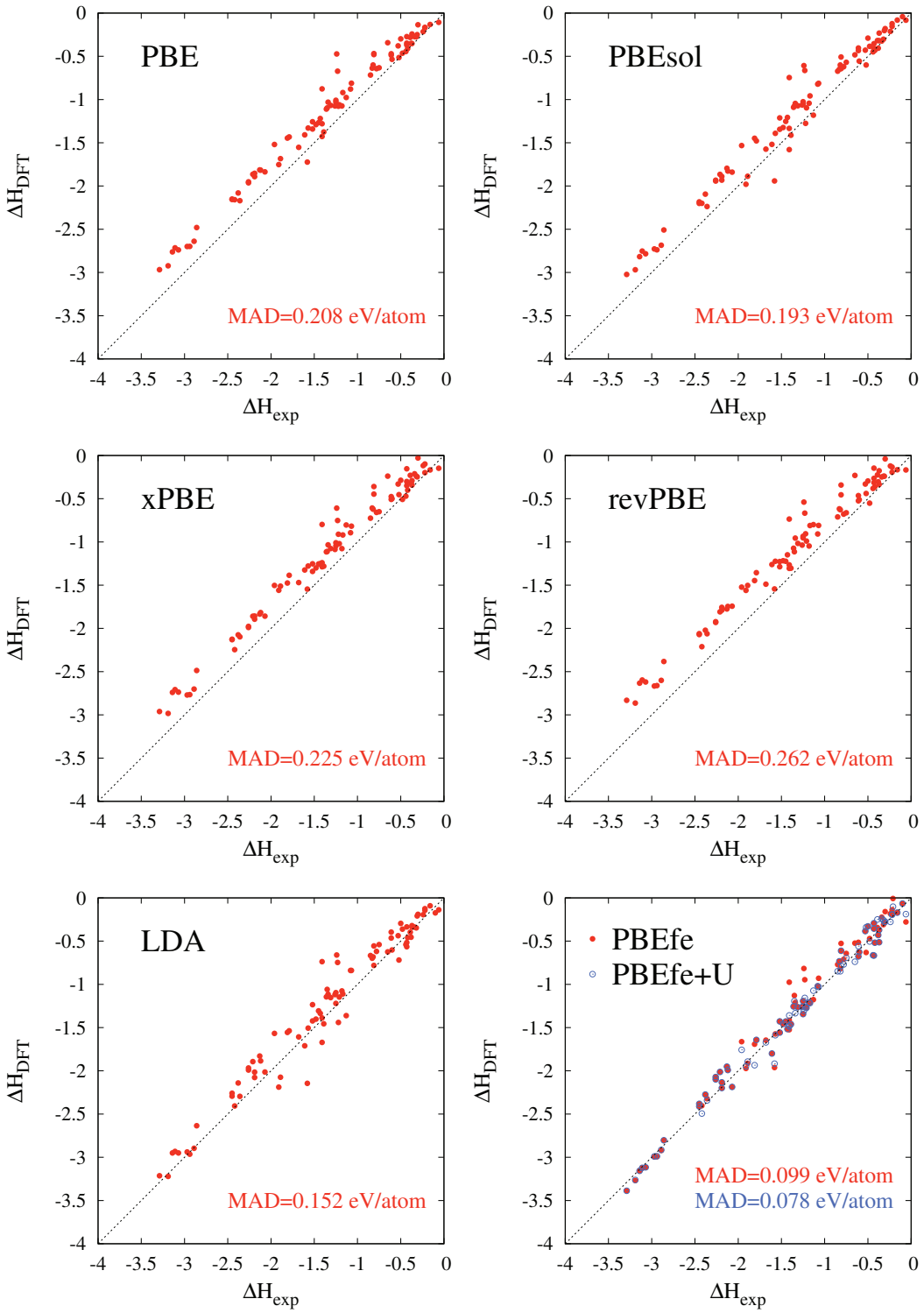


Fig. 4.3 Calculated versus experimental energies of formation of compounds in the optimization set for different exchange-correlation functionals. The corresponding MAD are indicated in the graphs.

found here. Other approaches like the use of FERE, can yield very small errors, as they include the ionic motion contribution on the fitted corrections of the total energies.

Table 4.3 Values of the Hubbard U in eV for the transition metals obtained to give the lowest energies of formation with the PBEfe functional.

Sc	3.0	La	0.5	Pd	0.0
Ni	3.0	Nb	0.0	Ti	2.7
Zn	5.3	V	6.0	Y	3.0
Co	1.3	Cd	3.3	Cr	1.0
Pt	0.0	Zr	3.0	Hg	5.7
Ag	6.0	Fe	2.0	Cu	3.0

Figure 4.4 shows the PBE and PBEfe enhancement factor $F_{\rm xc}(r_s, \zeta, s)$. In the range of interest for real systems, $0 \leq s \leq 3$ and $0 \leq r_s \leq 10$, it varies less than the enhancement factor of the PBE. The simultaneous increasing of μ and decreasing of κ makes the non-locality of PBEfe, i.e. the dependence with s, larger than in PBE for $s \sim 1$ and $1 \leq r_s \leq 10$, while as s increases, the dependence becomes smaller. Decreasing β, and increasing μ destroys the cancellation between the exchange and correlation terms for the small s limit ($s \to 0$). This makes the concave bump of $F_{\rm xc}$ more pronounced while $F_{\rm xc}(r_s, \zeta, s \sim 0)$ remains as in PBE and $F_{\rm xc}(r_s, \zeta, s \sim 3)$ varies slightly. If β were decreased according to the relation $\mu = \pi^2\beta/3$, $F_{\rm xc}$ would become convex as r_s increases. This is because the density for which the exchange stops being dominant would be smaller, decreasing the gradient correction for r_s lower than 10.

A minimization of only the two parameters of exchange was also performed (i.e., keeping the PBE correlation part). The lowest MAD was found of 0.155 eV/atom for $\mu = 0.255$ and $\kappa = 0.395$. For this case $F_{\rm xc}$ becomes almost constant for $r_s \sim 5$, being convex for $r_s > 5$ and concave for $r_s < 5$.

PBEfe is designed exclusively to obtain energies of formation or distances to the convex hull. It is however interesting to study what happens for other quantities that are also important for high-throughput investigations. One of this is the geometry. It is well known that PBE functional usually overestimates the lattice constants, while LDA underestimates them. This can be verified in Figure 4.5, where the green bars dominate the positive side while the blue ones prevail in the negative side. When it comes to the PBEfe functional, Figure 4.5 shows that the abundance of orange bars, corresponding to the PBEfe energies, is not significantly larger on one side than on the other. In fact approximately a 47% of the lattice parameters computed with PBEfe

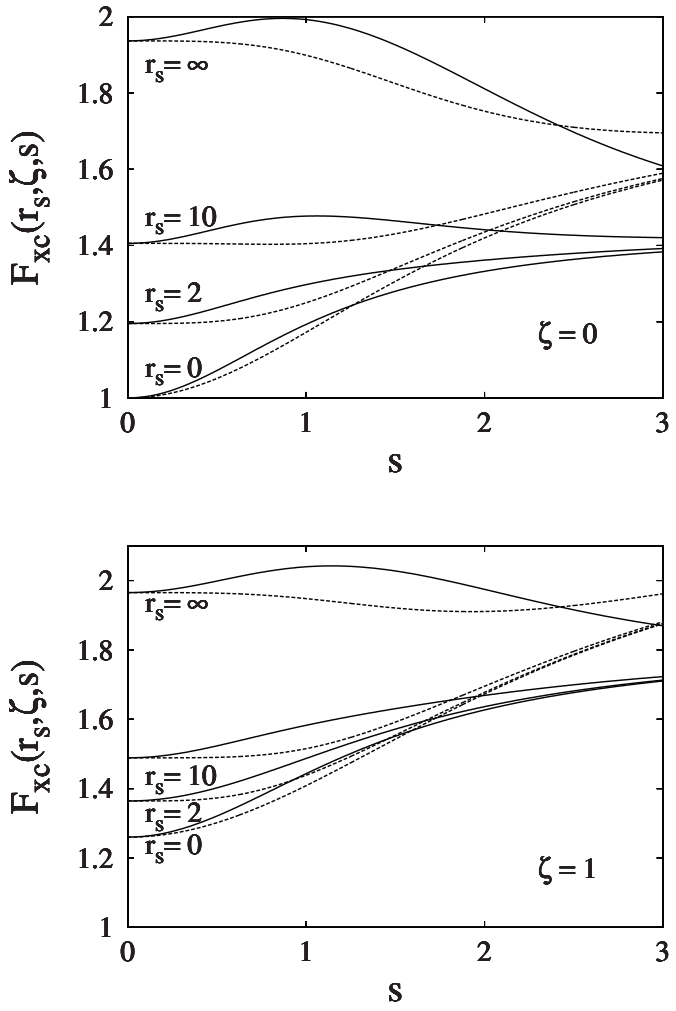


Fig. 4.4 Enhancement factors (Equation 4.7) for the spin-unpolarized ($\zeta = 0$) and the ferromagnetic spin-polarized ($\zeta = 1$) cases. Solid curves represent the fitted functional and the dashed ones represent the PBE functional.

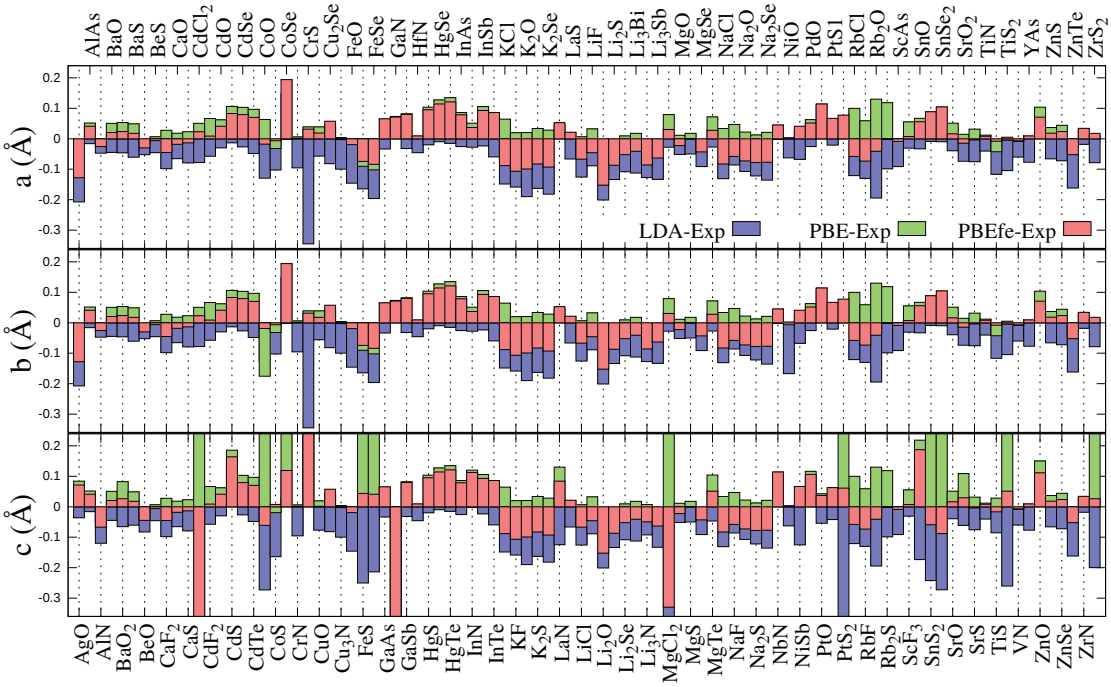


Fig. 4.5 Differences between the experimental and calculated lattice constants for PBE, PBEfe, and LDA functionals.

are lower than the experimental ones (that is considering all negative values, even the ones very close to experiment).

Figure 4.5 also shows that the PBEfe functional yields lattice constants, for most of the cases, lying between the values computed with LDA and PBE. The MAD of the values of the a and b cell parameters with respect to experiment is 0.051 Å for PBEfe, to be compared with 0.049 Å for the PBE and 0.080 Å for the LDA. For the c parameter the MAD is 0.068 Å for PBEfe, 0.150 Å for PBE and 0.119 Å for the LDA. We can therefore conclude that PBEfe is at least as good as the PBE for the geometries of solids, and often better when it comes to layered systems.

To verify the dependence of these errors on the choice of the optimization set, the PBEfe functional was tested in another set of compounds, different of the optimization set (results are summarized in Table C.4). This control set is formed by 104 binaries including 52 transition metal compounds and 52 alkali, alkali-earth or groups IIIA and IVA compounds and 33 ternary compounds (including one transition-metal sulfide, 7 oxides without transition metals, and 26 oxides with 1 or 2 transition metals).

The results can be seen in Figure 4.6 in comparison with PBE and LDA. It can be seen that the PBEfe functional corrects again to a large extent the systematic errors that affect the PBE energies of formation also for the control set. The conclusions are

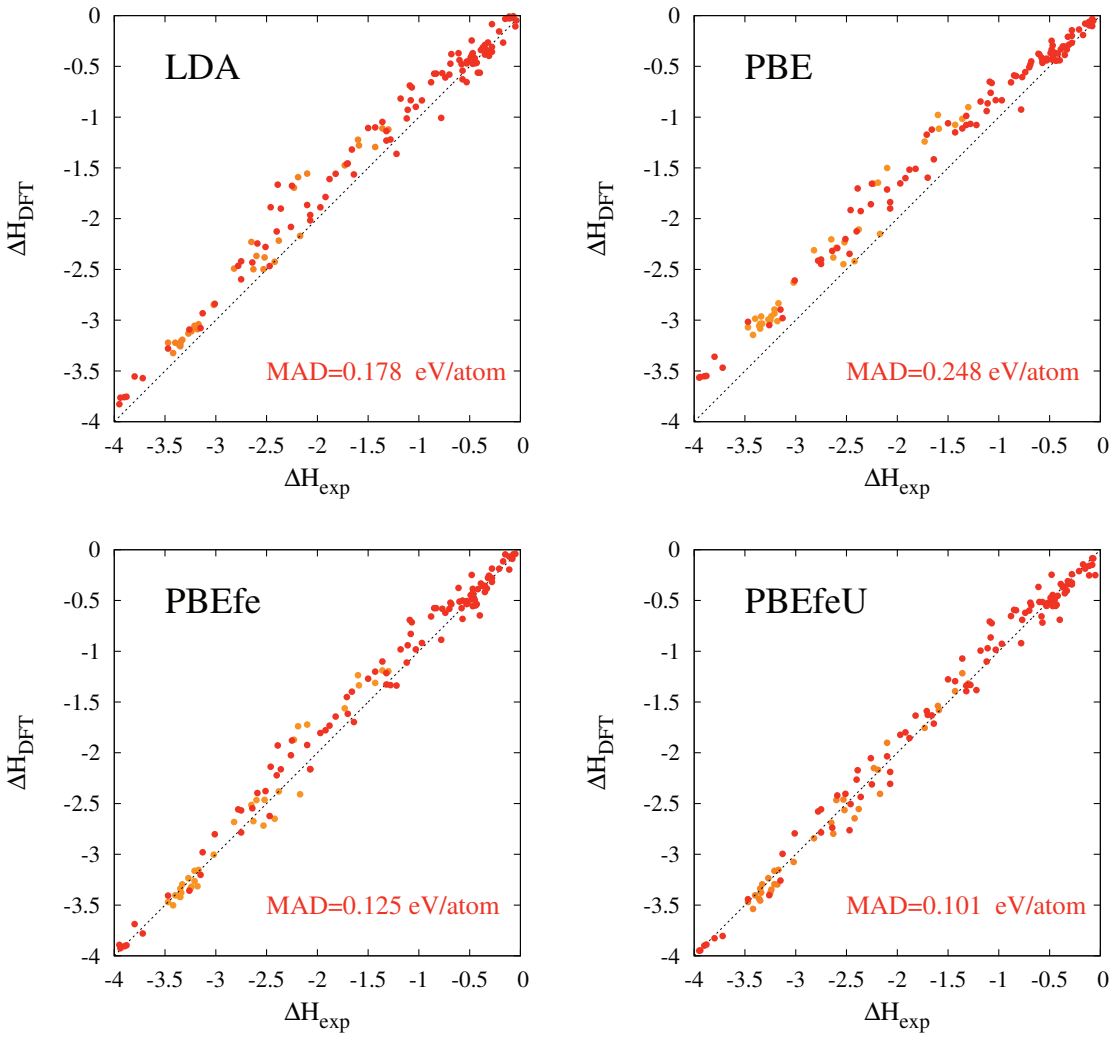


Fig. 4.6 Calculated versus experimental energies of formation in the control set for different exchange-correlation functionals. Orange dots correspond to ternary compounds. The MAD is indicated in each graph.

essentially the same as for the optimization set, i.e., the PBEfe MAD is substantially smaller than for the PBE (a 50% decrease for this set, and a 59% if a Hubbard U is used for transition metals).

In conclusion, we presented a new semi-empirical functional capable of reducing by 50% the error in the calculation of formation energies. This is achieved by fitting only *three parameters* of the PBE exchange-correlation functional to experimental formation energies. This means that this functional is trivially implemented in any code that already has the PBE functional, with no additional computational costs. The PBEfe functional can also be easily coupled to post-processing techniques to reduce the error

in the formation energies like FERE. Furthermore, the theoretical lattice constants are at least as good as the ones calculated with the PBE. For these reasons we think that our PBEfe functional is an excellent choice to search for new materials either using structure prediction or in high-throughput investigations.

Chapter 5

General Conclusions

This manuscript gathers several examples of developments and new applications of structure prediction and theoretical spectroscopy. We employed modern methods based on Density Functional Theory to define our first-principles approach to materials design. The variety of the systems considered here, ranging from intermetallic alloys and semiconductors to molecular solids, shows that these techniques are general and can be used in diverse domains of science.

The different topics have been organized in three chapters. Chapter 2 is devoted to the study of binary phase diagrams. We developed a methodology based on the Minima Hopping method for crystal structure prediction, to compute phase diagrams of binary compounds from first-principles in an easy and automatizable way. We first, applied this procedure to compute the phase diagrams of sodium-gold and lithium-aluminum binaries. We determined and characterized several new stable phases of these systems. Moreover, in both cases, we found all already experimentally-reported structures, validating the predictive power of the Minima Hopping method. We found interesting features like the increasing of the negative charge on gold with the sodium content, showing Bader charges ranging from -0.4 to -2.25. We also showed that the stiffness of the lithium-aluminum ordered alloys, with up to 60% of lithium, remains essentially equal to the one of aluminum while its bulk modulus decreases almost monotonically with the lithium content. We calculated the phase diagram of hydrogen-chlorine binaries and we found that the H_3 molecule can be stabilized in these compounds at moderate pressures of about 100 GPa, much lower than what predicted for pure hydrogen (2.1 TPa). This could open new opportunities to study this exotic molecule in an anvil cell in any laboratory.

In Chapter 3 we studied materials of interest for applications in photovoltaics. First, we considered the transparent conductive oxide $CuBO_2$, We have found a new stable crystal structure quite lower in energy than the previously proposed delafossite structure. We crosschecked our results with the fact that our structure prediction calculations gave the for $CuAlO_2$, $CuGaO_2$ and $CuInO_2$ the delafossite structure as the ground state, in agreement with experiment. This result is of big interest for

experimentalist due the technological relevance of this compound, which is a promising p-type transparent conductor, and not yest well studied.

Going an step further, we proposed then an efficient approach to search for new stable materials, combining standard prototype search and state-of-the-art global structure prediction methods. This method was applied to the discovery of new stable ABN_3 perovskite structures. From the nearly 2000 possible compositions of the form ABN_3, only two stable materials were known experimentally (and do not crystallize in the perovskite structure). We found 21 new compositions that have an energy below the known convex hull of thermodynamic stability. This, by itself, shows that modern computational techniques offer an effective tool to screen thousands of possible materials in a fraction of the time required by an equivalent experimental search. The unexpected discovery of stable nitride perovskites puts into question the previously accepted domain of stability of this crystal structure and opens the way for the experimental synthesis of exotic perovskites with new potential applications.

The last section of the chapter presents the study of the $Cu(In,Ga)S_2$ chalcopyrite alloys. Our calculations using a simple band alignment scheme, allowed us to reproduce the behavior of the X-ray absorption edges measured by collaborators. We found that the observed change in the band gap with changing alloy composition results from a spatial average of the nearly unchanged element-specific local states. Although the study was performed for the specific case of CIGS, we believe that our physical interpretation is also applicable to other compounds with different chemical composition.

Finally, while performing many calculations of total energies for varied systems, we realized that the state-of-the-art approximations do not yield very accurate formation energies, especially when we consider energy differences between dissimilar systems. We developed therefore an exchange-correlation functional based on the Perdew-Burke-Ernzerhof (PBE) approximation called PBEfe, which yields errors of the energies of formation smaller by a factor of 2 than the ones obtained with PBE, without increasing the computational time. This functional is specifically designed to be applied in high-throughput calculations. The optimization was done using a set of binary compounds and the functional was tested against a different set containing both binaries and ternary compounds, yielding a similar reduction of errors (of about a factor of 2). Moreover, these results were obtained without affecting the quality of the PBE geometries, which is an essential requirement to use this functional for of materials design.

We are confident that the PBE functional will be appreciated by the community: it improves significantly the quality of the energies of formation and it is easy to

implement as most of the codes already include a PBE functional. PBEfe has been included in the LIBXC [175] library of exchange-correlation functionals.

In the future, we plan to apply the Downhill Simplex Method that we used, to also optimize for determining accurate energies of formation, exchange-correlation functionals not derived from the PBE approximation. Moreover, we are working on the development of new techniques to reduce the error of the energies of formation based Artificial Neural Networks, which have already been successfully used to predict heat of formation of molecules.

List of Publications

Rafael Sarmiento-Pérez, Tiago F. T. Cerqueira, Irais Valencia-Jaime, Maximilian Amsler, Stefan Goedecker, Aldo H. Romero, Silvana Botti, and Miguel A. L. Marques. Novel phases of lithium-aluminum binaries from first-principles structural search. *J. Chem. Phys.*, 142, 024710, 2015

Sabine Körbel, David Kammerlander, **Rafael Sarmiento-Pérez** Claudio Attaccalite, Miguel A. L. Marques, and Silvana Botti. Optical properties of Cu-chalcogenide photovoltaic absorbers from self-consistent GW and the Bethe-Salpeter equation. *Phys. Rev. B*, 91, 075134, 2015

Rafael Sarmiento-Pérez, Silvana Botti, Claudia S. Schnohr, Iver Lauermann, Angel Rubio and Benjamin Johnson. Local versus global electronic properties of chalcopyrite alloys: X-ray absorption spectroscopy and ab initio calculations, *J. Appl. Phys.*, 116, 093703, 2014

Tiago F. T. Cerqueira, **Rafael Sarmiento-Pérez**, Fabio Trani, Maximilian Amsler, Stefan Goedecker, Miguel A. L. Marques, and Silvana Botti. The crystal structure of p-type transparent conductive oxide $CuBO_2$. *MRS Commun.*, 3, 157, 2013

Rafael Sarmiento-Pérez, Tiago F. T. Cerqueira, Irais Valencia-Jaime, Maximilian Amsler, Stefan Goedecker, Silvana Botti, Miguel A. L. Marques and Aldo H. Romero. Sodium–gold binaries: novel structures for ionic compounds from an ab initio structural search. *New J. Phys.*, 15, 115007, 2013

Rafael Sarmiento-Pérez, Tiago F. T. Cerqueira, Sabine Körbel, Silvana Botti and Miguel A. L. Marques. Prediction of stable nitride perovskites. *Accepted in Chem. Mater.*

Rafael Sarmiento-Pérez, Silvana Botti and Miguel A. L. Marques. Optimized exchange and correlation semi-local functional for the calculation of energies of formation. *Accepted in J. Chem. Theory Comput.*

Tiago F. T. Cerqueira, **Rafael Sarmiento-Pérez**, Maximilian Amsler, Fernando Nogeira, Silvana Botti and Miguel A L Marques. Materials Design on the fly. *Accepted in J. Chem. Theory Comput.*

Irais Valencia-Jaime, **Rafael Sarmiento-Pérez**, Maximilian Amsler, Stefan Goedecker, Silvana Botti, Miguel A. L. Marques and Aldo H. Romero. Crystal structures for lithium-silicon alloy predicted by minima hopping method. *Submitted to Phys. Rev. B*

References

[1] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Ann. Phys.*, vol. 389, p. 457, 1927.

[2] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, p. B864, 1964.

[3] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, p. A1133, 1965.

[4] G. E. Scuseria and V. N. Staroverov, "Chapter 24 - progress in the development of exchange-correlation functionals," in *Theory and Applications of Computational Chemistry* (C. E. Dykstra, F. Gernot, S. K. Kwang, and G. E. Scuseria, eds.), p. 669, Amsterdam: Elsevier, 2005.

[5] X. Xu and W. A. Goddard, "The extended Perdew-Burke-Ernzerhof functional with improved accuracy for thermodynamic and electronic properties of molecular systems," *J. Chem. Phys.*, vol. 121, p. 4068, 2004.

[6] J. P. Perdew and S. Kurth, "Chapter 1 - Density Functionals for Non-relativistic Coulomb Systems in the New Century," in *A primer in DFT* (M. A. L. M. Fernando Nogueira, Alberto Castro, ed.), p. 1, Springer-Verlag Berlin Heidelberg, 2003.

[7] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," *Phys. Rev. Lett.*, vol. 45, p. 566, 1980.

[8] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," *Phys. Rev. B*, vol. 45, p. 13244, 1992.

[9] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Can. J. Phys.*, vol. 58, p. 1200, 1980.

[10] J. P. Perdew and W. Yue, "Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation," *Phys. Rev. B*, vol. 33, p. 8800, 1986.

[11] M. Ernzerhof, J. P. Perdew, and K. Burke, "Density functionals: Where do they come from, why do they work?, in topics in current chemistry," in *Topics in Current Chemistry* (R. F. Nalewajski, ed.), vol. 180, p. 1, Springer, Berlin, 1996.

[12] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, p. 3865, 1996.

[13] E. H. Lieb and S. Oxford, "Improved lower bound on the indirect coulomb energy," *Int. J. Quantum Chem.*, vol. 19, p. 427, 1981.

[14] S. M. Woodley and R. Catlow, "Crystal structure prediction from first principles," *Nat. Mater.*, vol. 7, p. 937, 2008.

[15] E. Zurek and W. Grochala, "Predicting crystal structures and properties of matter under extreme conditions via quantum mechanics: the pressure is on," *Phys. Chem. Chem. Phys.*, vol. 17, p. 2917, 2015.

[16] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, "AFLOWLIB.ORG: A distributed materials properties repository from high-throughput *ab initio* calculations," *Comput. Mater. Sci.*, vol. 58, p. 227, 2012.

[17] J. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, "Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD)," *JOM*, vol. 65, p. 1501, 2013.

[18] C. J. Pickard and R. J. Needs, "Ab initio random structure searching," *J. Phys.: Condens. Matter*, vol. 23, p. 053201, 2011.

[19] J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, "Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling," *Phys. Rev. X*, vol. 4, p. 011019, 2014.

[20] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Neural Networks, 1995. Proceedings., IEEE International Conference on*, vol. 4, p. 1942, 1995.

[21] Y. Wang, J. Lv, L. Zhu, and Y. Ma, "Crystal structure prediction via particle-swarm optimization," *Phys. Rev. B*, vol. 82, p. 094116, 2010.

[22] Y. Wang and Y. Ma, "Perspective: Crystal structure prediction at high pressures," *J. Chem. Phys.*, vol. 140, p. 040901, 2014.

[23] S. Kirkpatrick, C. Gelatt, and M. Vecchi *Science*, vol. 220, p. 671.

[24] D. J. Wales and J. P. K. Doye, "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms," *J. Phy. Chem. A*, vol. 101, p. 5111, 1997.

[25] S. Goedecker, "Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems," *J. Chem. Phys.*, vol. 120, p. 9911, 2004.

[26] M. Amsler and S. Goedecker, "Crystal structure prediction using the minima hopping method," *J. Chem. Phys.*, vol. 133, p. 224104, 2010.

[27] S. Roy, S. Goedecker, and V. Hellmann, "Bell-Evans-Polanyi principle for molecular dynamics trajectories and its implications for global optimization," *Phys. Rev. E*, vol. 77, p. 056707, 2008.

[28] M. Sicher, S. Mohr, and S. Goedecker, "Efficient moves for global geometry optimization methods and their application to binary systems," *J. Chem. Phys.*, vol. 134, p. 044106, 2011.

[29] S. E. Schönborn, S. Goedecker, S. Roy, and A. R. Oganov, "The performance of minima hopping and evolutionary algorithms for cluster structure prediction," *J. Chem. Phys.*, vol. 130, p. 144108, 2009.

[30] K. Bao, S. Goedecker, K. Koga, F. Lançon, and A. Neelov, "Structure of large gold clusters obtained by global optimization using the minima hopping method," *Phys. Rev. B*, vol. 79, p. 041405, 2009.

[31] M. Amsler, J. A. Flores-Livas, L. Lehtovaara, F. Balima, S. A. Ghasemi, D. Machon, S. Pailhès, A. Willand, D. Caliste, S. Botti, A. San Miguel, S. Goedecker, and M. A. L. Marques, "Crystal structure of cold compressed graphite," *Phys. Rev. Lett.*, vol. 108, p. 065501, 2012.

[32] S. Botti, J. A. Flores-Livas, M. Amsler, S. Goedecker, and M. A. L. Marques, "Low-energy silicon allotropes with strong absorption in the visible for photovoltaic applications," *Phys. Rev. B*, vol. 86, p. 121204, 2012.

[33] J. A. Flores-Livas, M. Amsler, T. J. Lenosky, L. Lehtovaara, S. Botti, M. A. L. Marques, and S. Goedecker, "High-pressure structures of disilane and their superconducting properties," *Phys. Rev. Lett.*, vol. 108, p. 117004, 2012.

[34] M. Amsler, J. A. Flores-Livas, T. D. Huan, S. Botti, M. A. L. Marques, and S. Goedecker, "Novel structural motifs in low energy phases of $LiAlH_4$," *Phys. Rev. Lett.*, vol. 108, p. 205505, 2012.

[35] T. F. Cerqueira, R. Sarmiento-Pérez, F. Trani, M. Amsler, S. Goedecker, M. A. Marques, and S. Botti, "The crystal structure of p-type transparent conductive oxide $CuBO_2$," *MRS Commun.*, vol. 3, p. 157, 2013.

[36] R. Sarmiento-Pérez, T. F. Cerqueira, I. V.-J. e, M. Amsler, S. Goedecker, S. Botti, M. A. Marques, and A. H. Romero, "Sodium-gold binaries: novel structures for ionic compounds from an *ab initio* structural search," *New J. Phys.*, vol. 15, p. 115007, 2013.

[37] G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set," *Comput. Mat. Sci.*, vol. 6, p. 15, 1996.

[38] G. Kresse and J. Furthmüller, "Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set," *Phys. Rev. B*, vol. 54, p. 11169, 1996.

[39] X. Gonze, "First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm," *Phys. Rev. B*, vol. 55, p. 10337, 1997.

[40] X. Gonze and C. Lee, "Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory," *Phys. Rev. B*, vol. 55, p. 10355, 1997.

[41] H. Ikehata, N. Nagasako, T. Furuta, A. Fukumoto, K. Miwa, and T. Saito, "First-principles calculations for development of low elastic modulus Ti alloys," *Phys. Rev. B*, vol. 70, p. 174113, 2004.

[42] M. F. D. M. D. Raabe, B. Sander and J. Neugebauer, "Ab-initio simulation and experimental validation of β-titanium alloys," *Arxiv*, vol. abs/1207.0016, 2008.

[43] S. Sandlobes, Z. Pei, M. Friák, L.-F. Zhu, F. Wang, S. Zaefferer, D. Raabe, and J. Neugebauer, "Ductility improvement of mg alloys by solid solution: Ab initio modeling, synthesis and mechanical properties," *Acta Mater.*, vol. 70, p. 92, 2014.

[44] S. Hao, W. K. Liu, B. Moran, F. Vernerey, and G. B. Olson, "Multi-scale constitutive model and computational framework for the design of ultra-high strength, high toughness steels," *Comput. Methods Appl. Mech. Engrg.*, vol. 193, p. 1865, 2004. Multiple Scale Methods for Nanoscale Mechanics and Materials.

[45] F. P. Fehlner and S. M. Irving, "Sodium-gold alloy films," *J. of Appl. Phys.*, vol. 37, p. 3313, 1966.

[46] A. Pelton, "The Au-Na (gold-sodium) system," *Bull. Alloy Phase Diagr.*, vol. 7, p. 136, 1986. and references therein.

[47] U. Heiz, A. Vayloyan, E. Schumacher, C. Yeretzian, M. Stener, P. Gisdakis, and N. Rosch, "Na_xAu and Cs_xAu bimetal clusters: Finite size analogs of sodium-gold and cesium-gold compounds," *J. of Chem. Phys.*, vol. 105, p. 5574, 1996.

[48] K. Takemura and H. Fujihisa, "Na-Au intermetallic compounds formed under high pressure at room temperature," *Phys. Rev. B*, vol. 84, p. 014117, 2011.

[49] V. Smetana, Q. Lin, D. Pratt, A. Kreyssig, M. Ramazanoglu, J. D. Corbett, A. Goldman, and G. Miller, "A sodium-containing quasicrystal: Using gold to enhance sodium covalency in intermetallic compounds," *Angew. Chem. Int. Ed.*, vol. 51, p. 12699, 2012.

[50] E. Zintl, J. Goubeau, and W. Dullenkopf, "Metals and alloys. I. Salt-like compounds and intermetallic phases of sodium in liquid ammonia," *Z. Phys. Chem., Abt. A*, vol. 154, p. 1, 1931.

[51] T. B. Massalski, P. R. Subramanian, and L. Kacprzak, "Binary alloy phase diagrams, 2nd ed.," 1990.

[52] E. Havinga, H. Damsma, and P. Hokkeling, "Compounds and pseudo-binary alloys with the $CuAl_2$(C16)-type structure I. preparation and X-ray results," *J. Less-Common Met.*, vol. 27, p. 169, 1972.

[53] K. Lieser and H. Witte, "Löslichkeit von Wasserstoff in Legierungen IV. Diskussion der Ergebnisse," *Z. Metallkd*, vol. 43, p. 396, 1952.

[54] P. Villars and L. D. Calvert, "Pearson's handbook of crystallographic data for intermetallic phases, 2nd ed.," 1991.

[55] J. E. Ellis, "Adventures with substances containing metals in negative oxidation states," *Inorg. Chem.*, vol. 45, p. 3167, 2006.

[56] M. Jansen, "The chemistry of gold as an anion," *Chem. Soc. Rev.*, vol. 37, p. 1826, 2008.

[57] C. Koenig, N. E. Christensen, and J. Kollar, "Electronic properties of alkali-metal-gold compounds," *Phys. Rev. B*, vol. 29, p. 6481, 1984.

[58] T. Baruah, S. A. Blundell, and R. R. Zope, "Electronic and structural properties of small clusters of Na_nAu and Na_nAg (n=1–10)," *Phys. Rev. A*, vol. 64, p. 043202, 2001.

[59] Y.-C. Lin and D. Sundholm, "Computational studies of nonstoichiometric sodium auride clusters," *J. Phys. Chem. A*, vol. 116, p. 5119, 2012.

[60] R. King, "Metal cluster topology 19. beyond the auride ion: triangulated gold networks and ethane-like structural units in binary and ternary alkali metal gold intermetallics," *Inorg. Chim. Acta*, vol. 277, p. 202, 1998.

[61] B. Predel, "Au-Na (Gold-Sodium)," in *Ac-Au – Au-Zr* (O. Madelung, ed.), vol. 5a of *Landolt-Börnstein - Group IV Physical Chemistry*, p. 1, Springer Berlin Heidelberg, 1991.

[62] S. Baroni, P. Giannozzi, and A. Testa, "Green's-function approach to linear response in solids," *Phys. Rev. Lett.*, vol. 58, p. 1861, 1987.

[63] X. Gonze, "Perturbation expansion of variational principles at arbitrary order," *Phys. Rev. A*, vol. 52, p. 1086, 1995.

[64] X. Gonze, "Adiabatic density-functional perturbation theory," *Phys. Rev. A*, vol. 52, p. 1096, 1995.

[65] X. Gonze, G.-M. Rignanese, M. Verstraete, J.-M. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zérah, M. Mikami, P. Ghosez, M. Veithen, J.-Y. Raty, V. Olevano, F. Bruneval, L. Reining, R. Godby, G. Onida, D. Hamann, and D. Allan, "A brief introduction to the ABINIT software package," *Z. Kristallogr.*, vol. 220, p. 558, 2005.

[66] R. Bauer, A. Schmid, P. Pavone, and D. Strauch, "Electron-phonon coupling in the metallic elements Al, Au, Na, and Nb: A first-principles study," *Phys. Rev. B*, vol. 57, p. 11276, 1998.

[67] A. D. Corso, "Ab initio phonon dispersions of transition and noble metals: effects of the exchange and correlation functional," *J. Phys: Condens. Matter*, vol. 25, p. 145401, 2013.

[68] H. Qiao, T. Nohira, and Y. Ito, "Electrochemical formation of Au_2Na alloy and the characteristics of (Au_2Na+Au) reference electrode in a LiF – NaF – KF eutectic melt," *Electrochim. Acta*, vol. 47, p. 4543, 2002.

[69] B. Noble, S. Harris, and K. Dinsdale, "The elastic modulus of aluminium-lithium alloys," *J. Mater. Sci.*, vol. 17, p. 461, 1982.

[70] E. Lavernia and N. Grant, "Aluminium-lithium alloys," *J. Mater. Sci.*, vol. 22, p. 1521, 1987.

[71] E. Lavernia, T. Srivatsan, and F. Mohamed, "Strength, deformation, fracture behaviour and ductility of aluminium-lithium alloys," *J. Mater. Sci.*, vol. 25, p. 1137, 1990.

[72] J. Laverock, S. B. Dugdale, M. A. Alam, M. V. Roussenova, J. R. Wensley, J. Kwiatkowska, and N. Shiotani, "Fermi surface of an important nanosized metastable phase: Al_3li," *Phys. Rev. Lett.*, vol. 105, p. 236401, 2010.

[73] X.-Q. Guo, R. Podloucky, J.-h. Xu, and A. J. Freeman, "Cohesive, electronic, and structural properties of Al_3Li: An important metastable phase," *Phys. Rev. B*, vol. 41, p. 12432, 1990.

[74] X.-Q. Guo, R. Podloucky, and A. J. Freeman, "Phase stability and bonding characteristics of Li-rich Al-Li intermetallic compounds," *Phys. Rev. B*, vol. 42, p. 10912, 1990.

[75] Y. Hamon, T. Brousse, F. Jousse, P. Topart, P. Buvat, and D. M. Schleich, "Aluminum negative electrode in lithium ion batteries," *J. Power Sources*, vol. 97, p. 185, 2001.

[76] A. McAlister, "The Al-Li (aluminum-lithium) system," *Bull. Alloy Phase Diagr.*, vol. 3, p. 177, 1982.

[77] K. Puhakainen, M. Bostrom, T. L. Groy, and U. Haussermann, "A new phase in the system lithium-aluminum: Characterization of orthorhombic Li_2Al," *J. Solid State Chem.*, vol. 183, p. 2528, 2010.

[78] R. Podloucky, H. J. F. Jansen, X. Q. Guo, and A. J. Freeman, "First-principles electronic-structure approach for phase diagrams of binary alloys," *Phys. Rev. B*, vol. 37, p. 5478, 1988.

[79] A. Alam and D. D. Johnson, "Structural properties and relative stability of (meta)stable ordered, partially ordered, and disordered Al-Li alloy phases," *Phys. Rev. B*, vol. 85, p. 144202, 2012.

[80] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces," *Phys. Rev. Lett.*, vol. 100, p. 136406, 2008.

[81] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential," *J. Chem. Phys.*, vol. 118, p. 8207, 2003.

[82] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Erratum: "hybrid functionals based on a screened Coulomb potential" [j. chem. phys.118, 8207 (2003)]," *J. Chem. Phys.*, vol. 124, p. 219906, 2006.

[83] H. Okamoto, "Al-Li (Aluminum-Lithium)," *J. Phase Equil. Diffus.*, vol. 33, p. 500, 2012.

[84] T. Uesugi, Y. Takigawa, and K. Higashi, "Elastic Constants of AlLi from First Principles," *Mater. Trans.*, vol. 46, p. 1117, 2005.

[85] R. Hill, "The elastic behaviour of a crystalline aggregate," *Proc. Phys. Soc. Sect A*, vol. 65, p. 349, 1952.

[86] H. li Yu, X. hui Duan, Y. jun Ma, and M. Zeng, "First principles study of Al-Li intermetallic compounds," *Chin. J. Chem. Phys.*, vol. 25, p. 659, 2012.

[87] K. Kuriyama and S. Saito, "Elastic constants of single-crystal lithium indium," *Phys. Rev. B*, vol. 13, p. 1528, 1976.

[88] J. F. Thomas, "Third-order elastic constants of aluminum," *Phys. Rev.*, vol. 175, p. 955, 1968.

[89] J. Trivisonno and C. S. Smith, "Elastic constants of lithium-magnesium alloys," *Acta Metall.*, vol. 9, p. 1064, 1961.

[90] P. A. Korzhavyi, A. V. Ruban, S. I. Simak, and Y. K. Vekilov, "Electronic structure, thermal, and elastic properties of Al-Li random alloys," *Phys. Rev. B*, vol. 49, p. 14229, 1994.

[91] M. H. F. Sluiter, Y. Watanabe, D. d. Fontaine, and Y. Kawazoe, "First-principles calculation of the pressure dependence of phase equilibria in the Al-Li system," *Phys. Rev. B*, vol. 53, p. 6137, 1996.

[92] Y. Qi, L. G. Hector, C. James, and K. J. Kim, "Lithium concentration dependent elastic properties of battery electrode materials from first principles calculations," *J. Electrochem. Soc.*, vol. 161, p. F3010, 2014.

[93] G. Herzberg, "A spectrum of triatomic hydrogen," *J. Chem. Phys.*, vol. 70, p. 4806, 1979.

[94] M. Tashiro and S. Kato, "Quantum dynamics study on predissociation of H_3 Rydberg states: Importance of indirect mechanism," *J. Chem. Phys.*, vol. 117, p. 2053, 2002.

[95] C. Bordas, P. C. Cosby, and H. Helm, "Measurement of the lifetime of metastable triatomic hydrogen," *J. Chem. Phys.*, vol. 93, p. 6303, 1990.

[96] T. Shimanouchi, "Molecular vibrational frequencies," in *NIST Chemistry WebBook, NIST Standard Reference Database Number 69* (P. Linstrom and W. Mallard, eds.), National Institute of Standards and Technology, Gaithersburg MD, 20899: Springer, Berlin, 1996, 2015.

[97] M.-C. Chan, M. Okumura, and T. Oka, "Infrared spectrum of p-hydrogen crystals ionized by 3 mev electrons: cluster ions of hydrogen in condensed phase," *J. Phys. Chem. A*, vol. 104, p. 3775, 2000.

[98] M. Amsler, A. Mavel, R. Sarmiento-Pérez, S. Goedecker, S. Botti, and M. A. L. Marques, "Triatomic hydrogen in H-Cl binaries under moderate pressures," *Unpublished*.

[99] H. Liu, H. Wang, and Y. Ma, "Quasi-molecular and atomic phases of dense solid hydrogen," *J. Phys. Chem. C*, vol. 116, p. 9221, 2012.

[100] M. Snure and A. Tiwari, "$CuBO_2$: A p-type transparent oxide," *Appl. Phys. Lett.*, vol. 91, p. 092123, 2007.

[101] T. Minami, "New n-type transparent conducting oxides," *MRS Bull.*, vol. 25, p. 38, 2000.

[102] D. S. Ginley and C. Bright, "Transparent conducting oxides," *MRS Bull.*, vol. 25, p. 15, 2000.

[103] G. Hautier, A. Miglio, G. Ceder, G.-M. Rignanese, and X. Gonze, "Identification and design principles of low hole effective mass p-type transparent conducting oxides," *Nat. Commun.*, vol. 4, p. 3292, 2006.

[104] P. D. C. King and T. D. Veal, "Conductivity in transparent oxide semiconductors," *J. Phys: Condens. Matter*, vol. 23, p. 334214, 2011.

[105] H. Kawazoe, M. Yasukawa, H. Hyodo, M. Kurita, H. Yanagi, and H. Hosono, "p-type electrical conduction in transparent thin films of $CuAlO_2$," *Nature*, vol. 389, p. 939, 1997.

[106] D. O. Scanlon, A. Walsh, and G. W. Watson, "Understanding the p-type conduction properties of the transparent conducting oxide $CuBO_2$: A density functional theory analysis," *Chem. Mater.*, vol. 21, p. 4568, 2009.

[107] F. Trani, J. Vidal, S. Botti, and M. A. L. Marques, "Band structures of delafossite transparent conductive oxides from a self-consistent GW approach," *Phys. Rev. B*, vol. 82, p. 085115, 2010.

[108] S. Botti and M. A. L. Marques, "Strong renormalization of the electronic band gap due to lattice polarization in the GW formalism," *Phys. Rev. Lett.*, vol. 110, p. 226404, 2013.

[109] H. T. Stokes and D. M. Hatch, "FINDSYM: Program for identifying the space group symmetry of a crystal.," *J. Appl. Crystallogr.*, vol. 38, p. 237, 2005.

[110] B. U. Köhler and M. Jansen, "Darstellung und Strukturdaten von Delafossiten" $CuMO_2$ (M = Al, Ga, Sc, Y)," *Z. Anorg. Allg. Chem.*, vol. 543, p. 73, 1986.

[111] M. Ernzerhof and G. E. Scuseria, "Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional," *J. Chem. Phys.*, vol. 110, p. 5029, 1999.

[112] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study," *Phys. Rev. B*, vol. 57, p. 1505, 1998.

[113] Innovative Research and Products (iRAP) Inc. ET112: Piezoelectric Actuators and Motors-Types, Applications, new developments, Industry Structure and Global Markets (2010).

[114] M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Q. Wang, and C. W. Chu, "Superconductivity at 93 k in a new mixed-phase Y-Ba-Cu-O compound system at ambient pressure," *Phys. Rev. Lett.*, vol. 58, p. 908, 1987.

[115] M. A. Beno, L. Soderholm, D. W. Capone, II, D. G. Hinks, and J. D. Jorgensen, "Structure of the single-phase high-temperature superconductor $YBa_2Cu_3O_{(7-\delta)}$," *Appl. Phys. Lett.*, vol. 51, p. 57, 1987.

[116] A. P. Ramirez, "Colossal magnetoresistance," *J. Phys: Condens. Matter*, vol. 9, p. 8171, 1997.

[117] N. Hur, S. Park, P. A. Sharma, J. S. Ahn, S. Guha, and S.-W. Cheong, "Electric polarization reversal and memory in a multiferroic material induced by magnetic fields," *Nature*, vol. 429, p. 392, 2004.

[118] W. Eerenstein, N. D. Mathur, and J. F. Scott, "Multiferroic and magnetoelectric materials," *Nature*, vol. 442, p. 759, 2006.

[119] M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami, and H. J. Snaith, "Efficient hybrid solar cells based on meso-superstructured organometal halide perovskites," *Science*, vol. 338, p. 643, 2012.

[120] M. Liu, M. B. Johnston, and H. J. Snaith, "Efficient planar heterojunction perovskite solar cells by vapour deposition," *Nature*, vol. 501, p. 395, 2013.

[121] T. Okuda, K. Nakanishi, S. Miyasaka, and Y. Tokura, "Large thermoelectric response of metallic perovskites: $Sr_{1-x}La_xTiO_3$ ($0 <\sim x <\sim 0.1$)," *Phys. Rev. B*, vol. 63, p. 113104, 2001.

[122] H. Tanaka and M. Misono, "Advances in designing perovskite catalysts," *Curr. Opin. Solid State Mater. Sci.*, vol. 5, p. 381, 2001.

[123] Y. Nishihata, J. Mizuki, T. Akao, H. Tanaka, M. Uenishi, M. Kimura, T. Okamoto, and N. Hamada, "Self-regeneration of a pd-perovskite catalyst for automotive emissions control," *Nature*, vol. 418, p. 164, 2002.

[124] G. Xing, N. Mathews, S. S. Lim, N. Yantara, X. Liu, D. Sabba, M. Grätzel, S. Mhaisalkar, and T. C. Sum, "Low-temperature solution-processed wavelength-tunable perovskites for lasing," *Nat. Mater*, vol. 13, p. 476, 2014.

[125] F. Deschler, M. Price, S. Pathak, L. E. Klintberg, D.-D. Jarausch, R. Higler, S. Hüttner, T. Leijtens, S. D. Stranks, H. J. Snaith, M. Atatüre, R. T. Phillips, and R. H. Friend, "High photoluminescence efficiency and optically pumped lasing in solution-processed mixed halide perovskite semiconductors," *J. Phys. Chem. Lett.*, vol. 5, p. 1421, 2014.

[126] R. Caracas and R. E. Cohen, "Prediction of polar ordered oxynitride perovskites," *Appl. Phys. Lett.*, vol. 91, p. 092902, 2007.

[127] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. a. Persson, "The Materials Project: A materials genome approach to accelerating materials innovation," *APL Mater.*, vol. 1, p. 011002, 2013.

[128] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, "Python materials genomics (pymatgen): A robust, open-source python library for materials analysis," *Comput. Mat. Sci.*, vol. 68, p. 314, 2013.

[129] H. Glawe, A. Sanna, E. K. U. Gross, and M. A. L. Marques, "A modified Pettifor chemical scale for data mining," *(to be published).*

[130] D. Pettifor, "A chemical scale for crystal-structure maps," *Solid State Commun.*, vol. 51, p. 31, 1984.

[131] B. Johnson, J. Klaer, S. Merdes, M. Gorgoi, B. Höpfner, A. Vollmer, and I. Lauermann, "Limitations of near edge X-ray absorption fine structure as a tool for observing conduction bands in chalcopyrite solar cell heterojunctions," *J. Electron Spectrosc. Relat. Phenom.*, vol. 190, Part A, p. 42, 2013.

[132] C. S. Schnohr, H. Kämmer, C. Stephan, S. Schorr, T. Steinbach, and J. Rensberg, "Atomic-scale structure and band-gap bowing in Cu(In,Ga)Se_2," *Phys. Rev. B*, vol. 85, p. 245204, 2012.

[133] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, "Special quasirandom structures," *Phys. Rev. Lett.*, vol. 65, p. 353, 1990.

[134] A. van de Walle, P. Tiwary, M. de Jong, D. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, and Z.-K. Liu, "Efficient stochastic generation of special quasirandom structures," *Calphad*, vol. 42, p. 13, 2013.

[135] V. I. Anisimov, J. Zaanen, and O. K. Andersen, "Band theory and mott insulators: Hubbard u instead of stoner i," *Phys. Rev. B*, vol. 44, p. 943, 1991.

[136] J. Vidal, S. Botti, P. Olsson, J.-F. m. c. Guillemoles, and L. Reining, "Strong interplay between structure and electronic properties in CuIn(S,Se)$_2$: A first-principles study," *Phys. Rev. Lett.*, vol. 104, p. 056401, 2010.

[137] S. Botti, D. Kammerlander, and M. A. L. Marques, "Band structures of Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$ from many-body methods," *Appl. Phys. Lett.*, vol. 98, p. 241915, 2011.

[138] I. Aguilera, J. Vidal, P. Wahnón, L. Reining, and S. Botti, "First-principles study of the band structure and optical absorption of $CuGaS_2$," *Phys. Rev. B*, vol. 84, p. 085145, 2011.

[139] S. Eckner, H. Kämmer, T. Steinbach, M. Gnauck, A. Johannes, C. Stephan, S. Schorr, and C. S. Schnohr, "Atomic-scale structure, cation distribution, and bandgap bowing in Cu(In,Ga)S_2 and Cu(In,Ga)Se_2," *Appl. Phys. Lett.*, vol. 103, p. 081905, 2013.

[140] P. E. Blöchl, "Projector augmented-wave method," *Phys. Rev. B*, vol. 50, p. 17953, 1994.

[141] L. Gütay, D. Regesch, J. K. Larsen, Y. Aida, V. Depredurand, A. Redinger, S. Caneva, S. Schorr, C. Stephan, J. Vidal, S. Botti, and S. Siebentritt, "Feedback mechanism for the stability of the band gap of $CuInSe_2$," *Phys. Rev. B*, vol. 86, p. 045216, 2012.

[142] A. Schleife, F. Fuchs, C. Rödl, J. Furthmüller, and F. Bechstedt, "Branch-point energies and band discontinuities of III-nitrides and III-/II-oxides from quasiparticle band-structure calculations," *Appl. Phys. Lett.*, vol. 94, p. 012104, 2009.

[143] W. Mönch, "Branch-point energies and the band-structure lineup at Schottky contacts and heterostrucures," *J. Appl. Phys.*, vol. 109, p. 113724, 2011.

[144] J. Tersoff, "Theory of semiconductor heterojunctions: The role of quantum dipoles," *Phys. Rev. B*, vol. 30, p. 4874, 1984.

[145] B. Höffling, A. Schleife, C. Rödl, and F. Bechstedt, "Band discontinuities at Si-TCO interfaces from quasiparticle calculations: Comparison of two alignment approaches," *Phys. Rev. B*, vol. 85, p. 035305, 2012.

[146] G.Bunker, *Introduction to XAFS A Practical Guide to X-ray Absorption Fine Structure Spectroscopy* . Cambridge University Press, 2010.

[147] J. C. Mikkelsen and J. B. Boyce, "Atomic-scale structure of random solid solutions: Extended X-ray-absorption fine-structure study of $Ga_{1-x}In_xAs$," *Phys. Rev. Lett.*, vol. 49, p. 1412, 1982.

[148] C. S. Schnohr, L. L. Araujo, P. Kluth, D. J. Sprouster, G. J. Foran, and M. C. Ridgway, "Atomic-scale structure of $Ga_{1-x}In_xP$ alloys measured with extended X-ray absorption fine structure spectroscopy," *Phys. Rev. B*, vol. 78, p. 115201, 2008.

[149] C. S. Schnohr, "Binary and ternary random alloys," in *X-ray Absorption Spectroscopy of Semiconductors* (C. S. Schnohr and M. C. Ridgway, eds.), vol. 190 of *Springer Series in Optical Sciences*, p. 29, Springer Berlin Heidelberg, 2015.

[150] C. S. Schnohr, P. Kluth, L. L. Araujo, D. J. Sprouster, A. P. Byrne, G. J. Foran, and M. C. Ridgway, "Anisotropic vibrations in crystalline and amorphous InP," *Phys. Rev. B*, vol. 79, p. 195203, 2009.

[151] A. Balzarotti, N. Motta, A. Kisiel, M. Zimnal-Starnawska, M. T. Czyyk, and M. Podgórny, "Model of the local structure of random ternary alloys: Experiment versus theory," *Phys. Rev. B*, vol. 31, p. 7526, 1985.

[152] C. S. Schnohr, "Structural and electronic contributions to the bandgap bowing of (In,Ga)P alloys," *J. Phys. Condens. Matter*, vol. 24, p. 325802, 2012.

[153] S. Yamazoe, H. Kou, and T. Wada, "A structural study of Cu-In-Se compounds by X-ray absorption fine structure," *J. Mater. Res.*, vol. 26, p. 1504, 2011.

[154] S. Medling, Y. Lee, H. Zheng, J. F. Mitchell, J. W. Freeland, B. N. Harmon, and F. Bridges, "Evolution of magnetic oxygen states in Sr-doped $LaCoO_3$," *Phys. Rev. Lett.*, vol. 109, p. 157204, 2012.

[155] T. Saitoh, T. Mizokawa, A. Fujimori, M. Abbate, Y. Takeda, and M. Takano, "Electronic structure and magnetic states in $La_{1-x}Sr_xCoO_3$ studied by photoemission and X-ray-absorption spectroscopy," *Phys. Rev. B*, vol. 56, p. 1290, 1997.

[156] Y. Jiang, F. Bridges, N. Sundaram, D. P. Belanger, I. E. Anderson, J. F. Mitchell, and H. Zheng, "Study of the local distortions of the perovskite system $La_{1-x}Sr_xCoO_3$ $(0 \leq x \leq 0.35)$ using the extended X-ray absorption fine structure technique," *Phys. Rev. B*, vol. 80, p. 144423, 2009.

[157] L. Wang, T. Maxisch, and G. Ceder, "Oxidation energies of transition metal oxides within the GGA+U framework," *Phys. Rev. B*, vol. 73, p. 195107, 2006.

[158] R. Gautier, X. Zhang, L. Hu, L. Yu, Y. Lin, T. O. L. Sunde, D. Chon, and A. Poeppelmeier, Kenneth R. Zunger, "Prediction and accelerated laboratory discovery of previously unknown 18-electron ABX compounds," *Nat. Chem.*, vol. 7, p. 308, 2015.

[159] V. Stevanović, S. Lany, X. Zhang, and A. Zunger, "Correcting density functional theory for accurate predictions of compound enthalpies of formation: Fitted elemental-phase reference energies," *Phys. Rev. B*, vol. 85, p. 115104, 2012.

[160] K. K. Irikura, "Experimental vibrational zero-point energies: Diatomic molecules," *J. Phys. Chem. Ref. Data*, vol. 36, p. 389, 2007.

[161] A. Jain, G. Hautier, S. P. Ong, C. J. Moore, C. C. Fischer, K. A. Persson, and G. Ceder, "Formation enthalpies by mixing GGA and GGA + U calculations," *Phys. Rev. B*, vol. 84, p. 045115, 2011.

[162] Y. Wang and J. P. Perdew, "Spin scaling of the electron-gas correlation energy in the high-density limit," *Phys. Rev. B*, vol. 43, p. 8911, 1991.

[163] A. D. Becke, "Density functional calculations of molecular bond energies," *J. Chem. Phys.*, vol. 84, p. 4524, 1986.

[164] Y. Zhang and W. Yang, "Comment on "Generalized gradient approximation made simple"," *Phys. Rev. Lett.*, vol. 80, p. 890, 1998.

[165] V. Tognetti, P. Cortona, and C. Adamo, "A new parameter-free correlation functional based on an average atomic reduced density gradient analysis," *J. Chem. Phys.*, vol. 128, p. 034101, 2008.

[166] V. Tognetti, P. Cortona, and C. Adamo, "Increasing physical constraints and improving performances in a parameter-free GGA functional," *Chem. Phys. Lett.*, vol. 460, p. 536, 2008.

[167] G. K. H. Madsen, "Functional form of the generalized gradient approximation for exchange: The PBE α functional," *Phys. Rev. B*, vol. 75, p. 195108, 2007.

[168] J. Klimeš, D. R. Bowler, and A. Michaelides, "Chemical accuracy for the van der Waals density functional," *J. Phys: Condend. Matter*, vol. 22, p. 022201, 2010.

[169] A. Ruzsinszky, G. I. Csonka, and G. E. Scuseria, "Regularized gradient expansion for atoms, molecules, and solids," *J. Chem. Theory Comput*, vol. 5, p. 763, 2009.

[170] E. Fabiano, L. A. Constantin, and F. Della Sala, "Generalized gradient approximation bridging the rapidly and slowly varying density regimes: A PBE-like functional for hybrid interfaces," *Phys. Rev. B*, vol. 82, p. 113104, 2010.

[171] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids," *Phys. Rev. Lett.*, vol. 91, p. 146401, 2003.

[172] P. Villars and J. Daams, "Atomic-environment classification of the chemical elements," *J. Alloys Compd.*, vol. 197, p. 177, 1993. Proceeding of the Workshop on Regularities, Classifications and Prediction of Advanced Materials.

[173] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, p. 308, 1965.

[174] S. Lany, "Semiconductor thermochemistry in density functional calculations," *Phys. Rev. B*, vol. 78, p. 245207, 2008.

[175] M. A. Marques, M. J. Oliveira, and T. Burnus, "Libxc: A library of exchange and correlation functionals for density functional theory," *Comput. Phys. Commun.*, vol. 183, p. 2272, 2012.

Appendix A

Sodium-Gold Alloys Crystal Structure Information

$NaAu_2$

Space Group: $F\,4_1/d\,\bar{3}\,2/m$ (Origin choice 1) (227)
$a = 7.9473$ Å $b = 7.9473$ Å $c = 7.9473$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Na at Wyckoff position b
Au at Wyckoff position c

NaAu

Space Group: $R\,\bar{3}\,2/m$ (Hexagonal axes) (166)
$a = 5.6542$ Å $b = 5.6542$ Å $c = 14.3409$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 120.00$

Na at Wyckoff position a
Na at Wyckoff position c, z = 0.3863
Au at Wyckoff position d

Na_3Au_2

Space Group: $C\,1\,2/c\,1$ (15)
$a = 7.8593$ Å $b = 8.5125$ Å $c = 7.0438$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Na at Wyckoff position e, y = 0.3865
Na at Wyckoff position f, x = 0.3154, y = 0.4312, z = -0.0106
Au at Wyckoff position f, x = 0.1323, y = 0.2701, z = -0.3655

Na_2Au

Space Group: $I\,4/m\,2/c\,2/m$ (140)
$a = 7.3823$ Å $b = 7.3823$ Å $c = 5.6288$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Na at Wyckoff position h, x = 0.3285
Au at Wyckoff position a

Na_3Au

Space Group: $P\,1\,2_1/m\,1$ (11)
$a = 5.7324$ Å $b = 9.2625$ Å $c = 4.5781$ Å
$\alpha = 90.00$ $\beta = 113.29$ $\gamma = 90.00$

Na at Wyckoff position f, x = -0.1950, y = -0.4363, z = 0.1536
Na at Wyckoff position e, x = -0.4001, z = -0.4508
Au at Wyckoff position e, x = -0.0053, z = 0.2464

Na_5Au

Space Group: $P\,1\,2_1/m\,1$ (125)
$a = 7.6310$ Å $b = 7.6310$ Å $c = 6.2395$ Å
$\alpha = 90.00$ $\beta = 113.29$ $\gamma = 90.00$

Na at Wyckoff position m, x = 0.1650, z = 0.2858
Na at Wyckoff position a
Au at Wyckoff position b

Appendix B

Lithium-Aluminum Alloys Crystal Structure Information

Li

Space Group $R\ \bar{3}\ m$ (166)
$a = 3.0672$ Å $b = 3.0672$ Å $c = 22.0849$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 120.00$

Li at Wyckoff position a
Li at Wyckoff position c, $z = 0.22223$
Li at Wyckoff position c, $z = 0.77777$

Li

Space Group $I\ m\ \bar{3}\ m$ (229)
$a = 3.4199$ Å $a = 3.4199$ Å $a = 3.4199$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Li at Wyckoff position a

Li

Space Group $F\ m\ \bar{3}\ m$ (225)
$a = 3.0429$ Å $a = 3.0429$ Å $a = 3.0429$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Li at Wyckoff position a

Al

Space Group $F4/m\ \bar{3}\ 2/m$ (255)
$a = 4.0412$ Å $b = 4.0412$ Å $c = 4.0412$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Al at Wyckoff position a

LiAl

Space Group $F\ 4_1/d\ \bar{3}\ 2/m$ (origin choice 1) (227)
$a = 6.3482$ Å $b = 6.3482$ Å $c = 6.3482$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Al at Wyckoff position a
Li at Wyckoff position b

$LiAl_3$

Space Group $P\ 4/m\ \bar{3}\ 2/m$ (221)
$a = 4.0245$ Å $b = 4.0245$ Å $c = 4.0245$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Al at Wyckoff position c
Li at Wyckoff position a

Li_2Al (Hexagonal)

Space Group $P\ 6/m\ 2/m\ 2/m$ (191)
$a = 4.5411$ Å $b = 4.5411$ Å $c = 2.6674$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 120.00$

Al at Wyckoff position b
Li at Wyckoff position c

Li_2Al (Orthorhombic)

Space Group $Cmcm$ (63)
$a = 4.5923$ Å $b = 9.5959$ Å $c = 4.4483$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 90.00$

Al at Wyckoff position c, y = 0.07673
Al at Wyckoff position c, y = -0.24257
Li at Wyckoff position c, y = 0.41064

Li_3Al

Space Group P 1 21/m 1 (11)
$a = 5.4198$ Å $b = 4.4599$ Å $c = 5.6365$ Å
$\alpha = 90.00$ $\beta = 100.03$ $\gamma = 90.00$

Al at Wyckoff position e, x = -0.3824, z = 0.4343
Li at Wyckoff position e, x = -0.3756, z = -0.0783
Li at Wyckoff position e, x = 0.1189, z = 0.2308
Li at Wyckoff position e, x = 0.1387, z = -0.3068

Li_3Al_2

Space Group $R\ \bar{3}\ 2/m$ (hexagonal axes)
$a = 4.4464$ Å $b = 4.4464$ Å $c = 14.0727$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 120.00$

Al at Wyckoff position c, z = -0.1981
Li at Wyckoff position a
Li at Wyckoff position c, z = 0.4029

Li_7Al_2

Space Group $R\ \bar{3}\ 2/m$ (hexagonal axes) (166)
$a = 7.9320$ Å $b = 7.9320$ Å $c = 8.2264$ Å
$\alpha = 90.00$ $\beta = 90.00$ $\gamma = 120.00$

Al at Wyckoff position c, z = 0.1644
Li at Wyckoff position f, x = -0.3288
Li at Wyckoff position b

Li_9Al

Space Group P-1 (2)
$a = 5.0759$ Å $b = 5.0925$ Å $c = 7.9362$ Å
$\alpha = 83.45$ $\beta = 71.71$ $\gamma = 70.30$

Al at Wyckoff position d
Li at Wyckoff position c
Li at Wyckoff position i, x = 0.4088, y = 0.4992, z = 0.1832
Li at Wyckoff position i, x = 0.3104, y = 0.0000, z = 0.3790
Li at Wyckoff position i, x = -0.2003, y = -0.5000, z = 0.4010
Li at Wyckoff position i, x = 0.0928, y = 0.0003, z = -0.1873

Li_9Al_4

Space Group $A\ 1\ 2/m\ 1$ (12)
$a = 5.3670$ Å $b = 4.4529$ Å $c = 18.6329$ Å
$\alpha = 90.00$ $\beta = 105.42$ $\gamma = 90.42$

Al at Wyckoff position i, x = 0.2825, z = -0.1512
Al at Wyckoff position i, x = -0.4324, z = 0.3862
Li at Wyckoff position i, x = -0.3425, z = -0.2330
Li at Wyckoff position i, x = -0.1379, z = -0.0848
Li at Wyckoff position i, x = -0.0276, z = 0.3079
Li at Wyckoff position i, x = -0.1755, z = -0.4565
Li at Wyckoff position d

Appendix C

Supplementary Tables for Chapter 3

Table C.1 Experimental and calculated lattice constants of the compounds in the optimization set.

Comp.	Space group	Exp[127] a	Exp[127] b	Exp[127] c	PBEfe a	PBEfe b	PBEfe c
AgO	$Cccm$	3.404	3.404	5.495	3.2760	3.2760	5.5666
AlAs	$F\bar{4}3m$	4.002	4.002	4.002	4.0430	4.0430	4.0430
AlN	$P6_3mc$	3.139	3.139	5.066	3.1135	3.1135	4.9990
BaO	$Fm\bar{3}m$	3.917	3.917	3.917	3.9373	3.9373	3.9373
BaO_2	$I4/mmm$	3.811	3.811	4.347	3.8349	3.8349	4.3734
BaS	$Fm\bar{3}m$	4.517	4.517	4.517	4.5351	4.5351	4.5351
BeO	$P6_3mc$	2.725	2.725	4.424	2.6953	2.6953	4.3791
BeS	$F\bar{4}3m$	3.440	3.440	3.440	3.4339	3.4339	3.4339
CaF_2	$Fm\bar{3}m$	3.868	3.868	3.868	3.8227	3.8227	3.8227
CaO	$Fm\bar{3}m$	3.402	3.402	3.402	3.3839	3.3839	3.3839
CaS	$Fm\bar{3}m$	4.018	4.018	4.018	4.0046	4.0046	4.0046
$CdCl_2$	$R\bar{3}m$	3.85	3.85	17.46	3.8732	3.8732	17.0854
CdF_2	$Fm\bar{3}m$	3.813	3.813	3.813	3.8217	3.8217	3.8217
CdO	$Fm\bar{3}m$	3.320	3.320	3.320	3.3610	3.3610	3.3610
CdS	$P6_3mc$	4.091	4.091	6.641	4.1742	4.1742	6.8050
CdSe	$F\bar{4}3m$	4.278	4.278	4.278	4.3572	4.3572	4.3572
CdTe	$F\bar{4}3m$	4.580	4.580	4.580	4.6503	4.6503	4.6503
CoO	$F\bar{4}3m$	3.190	3.190	5.439	3.1721	3.1710	5.3781
CoS	$P6_3/mmc$	3.380	3.380	5.150	3.3735	3.3735	5.1575
CoSe	$P4/nmm$	3.530	3.530	5.250	3.7246	3.7246	5.3691
CrN	$Fm\bar{4}m$	2.924	2.924	2.924	2.9256	2.9256	2.9256
CrS	$P6_3/mmc$	3.439	3.439	5.324	3.4704	3.4704	5.5896
CuO	$P4_2/mmc$	2.902	2.902	5.130	2.9208	2.9197	5.1316

Table C.1 Experimental and calculated lattice constants of the compounds in the optimization set.

	Space	Exp[127]			PBEfe		
Comp.	group	a	b	c	a	b	c
Cu_2Se	$Fm\bar{3}m$	4.075	4.075	4.075	4.1326	4.1326	4.1326
Cu_3N	$Pm\bar{3}m$	3.813	3.813	3.813	3.8099	3.8099	3.8099
FeO	$I4/mmm$	4.326	4.326	4.326	4.3066	4.3066	4.3066
FeS	$P4/nmm$	3.679	3.679	5.047	3.6048	3.6048	5.0907
FeSe	$P4/nmm$	3.779	3.779	5.511	3.6950	3.6950	5.5524
GaAs	$F\bar{4}3m$	3.997	3.997	3.997	4.0628	4.0628	4.0628
GaN	$P6_3mc$	3.142	3.142	5.720	3.2132	3.2132	5.2413
GaSb	$F\bar{4}3m$	4.310	4.310	4.310	4.3896	4.3896	4.3896
HfN	$Fm\bar{3}m$	3.200	3.200	3.200	3.2092	3.2092	3.2092
HgS	$F\bar{4}3m$	4.137	4.137	4.137	4.2321	4.2321	4.2321
HgSe	$F\bar{4}3m$	4.303	4.303	4.303	4.4176	4.4176	4.4176
HgTe	$F\bar{4}3m$	4.567	4.567	4.567	4.6881	4.6881	4.6881
InAs	$F\bar{4}3m$	4.285	4.285	4.285	4.3638	4.3638	4.3638
InN	$P6_3mc$	3.533	3.533	5.671	3.5702	3.5702	5.7839
InSb	$F\bar{4}3m$	4.580	4.580	4.580	4.6732	4.6732	4.6732
InTe	$Fm\bar{3}m$	4.368	4.368	4.368	4.4540	4.4540	4.4540
KCl	$Fm\bar{3}m$	4.440	4.440	4.440	4.3518	4.3518	4.3518
KF	$Fm\bar{3}m$	3.804	3.804	3.804	3.6974	3.6974	3.6974
K_2O	$Fm\bar{3}m$	4.560	4.560	4.560	4.4609	4.4609	4.4609
K_2S	$Fm\bar{3}m$	5.238	5.238	5.238	5.1551	5.1551	5.1551
K_2Se	$Fm\bar{3}m$	5.438	5.438	5.438	5.3449	5.3449	5.3449
LaN	$P6_3mc$	4.084	4.084	5.840	4.1369	4.1370	5.9240
LaS	$Fm\bar{3}m$	4.144	4.144	4.144	4.1655	4.1655	4.1655
LiCl	$Fm\bar{3}m$	3.637	3.637	3.637	3.5701	3.5701	3.5701
LiF	$Fm\bar{3}m$	2.850	2.850	2.850	2.8041	2.8041	2.8041
Li_2O	$Fm\bar{3}m$	3.382	3.382	3.382	3.2299	3.2299	3.2299
Li_2S	$Fm\bar{3}m$	4.073	4.073	4.073	3.9861	3.9861	3.9861
Li_2Se	$Fm\bar{3}m$	4.246	4.246	4.246	4.1941	4.1941	4.1941
Li_3Bi	$Fm\bar{3}m$	4.743	4.743	4.743	4.7019	4.7019	4.7019
Li_3N	$P6/mmm$	3.672	3.672	3.881	3.5859	3.5859	3.8318
Li_3Sb	$Fm\bar{3}m$	4.648	4.648	4.648	4.5850	4.5850	4.5850

Table C.1 Experimental and calculated lattice constants of the compounds in the optimization set.

Comp.	Space group	Exp[127] a	Exp[127] b	Exp[127] c	PBEfe a	PBEfe b	PBEfe c
$MgCl_2$	$R\bar{3}m$	3.596	3.596	17.589	3.6265	3.6265	17.2592
MgO	$Fm\bar{3}m$	2.998	2.998	2.998	2.9754	2.9754	2.9754
MgS	$Fm\bar{3}m$	3.678	3.678	3.678	3.6776	3.6776	3.6776
MgSe	$F\bar{4}3m$	4.246	4.246	4.246	4.2026	4.2026	4.2026
MgTe	$P6_3mc$	4.540	4.540	7.395	4.5680	4.5680	7.4461
NaCl	$Fm\bar{3}m$	3.988	3.988	3.988	3.9052	3.9052	3.9052
NaF	$Fm\bar{3}m$	3.266	3.266	3.266	3.2080	3.2080	3.2080
Na_2O	$Fm\bar{3}m$	3.932	3.932	3.932	3.8590	3.8590	3.8590
Na_2S	$Fm\bar{3}m$	4.629	4.629	4.629	4.5507	4.5507	4.5507
Na_2Se	$Fm\bar{3}m$	4.826	4.826	4.826	4.7491	4.7491	4.7491
NbN	$P\bar{6}m2$	2.940	2.940	2.790	2.9855	2.9855	2.9043
NiO	$Fm\bar{3}m$	2.944	5.088	2.944	2.9473	5.0815	2.9473
NiSb	$P6_3/mmc$	3.946	3.946	5.148	3.9871	3.9871	5.2148
PdO	$P4_2/mmc$	3.036	3.036	5.327	3.0877	3.0877	5.4333
PtO	$P4_2/mmc$	3.040	3.040	5.340	3.1544	3.1544	5.3769
PtS	$P4_2/mmc$	3.470	3.470	6.110	3.5369	3.5369	6.1729
PtS_2	$P\bar{3}m1$	3.537	3.537	5.019	3.6149	3.6149	5.0802
RbCl	$Fm\bar{3}m$	4.632	4.632	4.632	4.5738	4.5738	4.5738
RbF	$Fm\bar{3}m$	3.988	3.988	3.988	3.9146	3.9146	3.9146
Rb_2O	$R\bar{3}m$	6.74	6.74	6.74	6.6988	6.6988	6.6988
Rb_2S	$Fm\bar{3}m$	5.420	5.420	5.420	5.4186	5.4186	5.4186
ScAs	$Fm\bar{3}m$	3.881	3.881	3.881	3.8725	3.8725	3.8725
ScF_3	$Pm\bar{3}m$	4.011	4.011	4.011	4.0181	4.0181	4.0181
SnO	$P4/nmm$	3.800	3.800	4.810	3.8561	3.8561	4.9977
SnS_2	$P\bar{3}m1$	3.640	3.640	5.900	3.7290	3.7290	5.8409
$SnSe_2$	$P\bar{3}m1$	3.811	3.811	6.141	3.9156	3.9156	6.0528
SrO	$Fm\bar{3}m$	3.629	3.629	3.629	3.6451	3.6451	3.6451
SrO_2	$I4/mmm$	3.557	3.557	4.133	3.5426	3.5426	4.1625
SrS	$Fm\bar{3}m$	4.256	4.256	4.256	4.2520	4.2520	4.2520
TiN	$Fm\bar{3}m$	2.995	2.995	2.995	3.0032	3.0032	3.0032
TiS	$P6_3/mmc$	3.300	3.300	6.440	3.2917	3.2916	6.4235

Table C.1 Experimental and calculated lattice constants of the compounds in the optimization set.

	Space	Exp[127]			PBEfe		
Comp.	group	a	b	c	a	b	c
TiS_2	$P\bar{3}m1$	3.412	3.412	5.695	3.4171	3.4171	5.7463
VN	$Fm\bar{3}m$	2.925	2.925	2.925	2.9185	2.9185	2.9185
YAs	$Fm\bar{3}m$	4.112	4.112	4.112	4.1212	4.1212	4.1212
ZnO	$P6_3mc$	3.184	3.184	5.155	3.2547	3.2547	5.2666
ZnS	$P3m1$	3.813	3.813	3.813	3.8313	3.8313	3.8313
ZnSe	$F\bar{4}3m$	4.012	4.012	4.012	4.0356	4.0356	4.0356
ZnTe	$F\bar{4}3m$	4.402	4.402	4.402	4.3502	4.3502	4.3502
ZrN	$Fm\bar{3}m$	3.235	3.235	3.235	3.2691	3.2691	3.2691
ZrS_2	$P\bar{3}m1$	3.680	3.680	5.850	3.6977	3.6977	5.8765

Table C.2 Experimental and calculated lattice constants of the compounds in the control set.

Comp.	Space group	Exp[127] a	Exp[127] b	Exp[127] c	PBEfe a	PBEfe b	PBEfe c
Ag_2O	$Pn\bar{3}m$	4.723	4.723	4.723	4.7942	4.7942	4.7942
Ag_2S	$P2_1$	4.200	6.860	8.043	4.0073	5.8436	6.9808
Ag_2Se	$P2_12_12_1$	4.337	7.070	7.773	4.5332	7.0766	7.7835
$AlCl_3$	$C2/m$	5.910	5.910	6.148	5.9501	5.9501	6.0634
AlF_3	$R\bar{3}c$	5.030	5.030	5.030	5.0189	5.0189	5.0189
Al_2O_3	$R\bar{3}c$	5.128	5.128	5.128	5.1502	5.1502	5.1502
Al_2S_3	$P6_1$	6.491	6.491	17.169	6.5301	6.5302	17.275
Al_2Se_3	Cc	6.730	6.740	7.320	6.7703	6.7979	7.4215
Al_2Te_3	$P2_1/c$	7.181	12.848	14.167	7.2761	12.9880	14.4066
Be_3N_2	$Ia3$	7.054	7.054	7.054	7.0230	7.0230	7.0230
$CaCl_2$	$P4_2/mnm$	4.204	6.383	6.383	4.1230	6.3351	6.3351
CdSb	$Pbca$	6.471	8.253	8.526	6.6357	8.3637	8.7085
Cd_3As_2	$P4_2/nmc$	8.945	8.945	12.650	9.1353	9.1353	12.8584
Cd_3N_2	$Ia\bar{3}$	9.378	9.378	9.378	9.5031	9.5031	9.5031
CoF_2	$P4_2/mnm$	3.170	4.694	4.694	3.1988	4.6133	4.6133
CoF_3	$R\bar{3}c$	5.281	5.281	5.281	5.1217	5.1189	5.1189
$CoSb_3$	$Im\bar{3}$	7.864	7.864	7.864	7.9413	7.9413	7.9413
Co_3O_4	$Fd\bar{3}m$	5.736	5.736	5.736	5.7121	5.7121	5.7121
Co_3S_4	$Fd\bar{3}m$	6.651	6.651	6.651	6.6016	6.6016	6.6016
CrO_2	$P4_2/mnm$	2.897	4.371	4.371	2.9186	4.4306	4.4306
Cr_2O_3	$R\bar{3}c$	5.396	5.396	5.396	5.3628	5.3628	5.3628
CuF_2	$P2_1/c$	3.294	4.568	4.614	3.1383	4.6144	4.6062
CuS	$P6_3/mmc$	3.760	3.782	16.235	3.7751	3.8119	16.4588
Cu_2O	$Pn\bar{3}m$	4.268	4.268	4.268	4.2676	4.2676	4.2676
Cu_2Sb	$P4/nmm$	4.002	4.002	6.104	4.0651	4.0651	6.1256
Cu_2Te	$P6/mmm$	4.237	4.237	7.274	4.2766	4.2766	8.5022
Cu_3Sb	$Pmnm$	4.353	4.768	5.504	4.1829	5.6269	5.3493
CuSe	$Cmcm$	3.952	4.003	17.235	3.9657	4.0435	17.3670
FeF_2	$P4_2/mnm$	3.360	4.830	4.830	2.9330	4.9275	4.9275
Fe_2O_3	$R\bar{3}c$	5.431	5.431	5.431	4.7667	5.2173	4.7667
Fe_3O_4	$Fd\bar{3}m$	5.951	5.951	5.948	5.7536	5.7536	5.7537

Table C.2 Experimental and calculated lattice constants of the compounds in the optimization set.

	Space	Exp[127]			PBEfe		
Comp.	group	a	b	c	a	b	c
$GaCl_3$	$C2/m$	6.855	6.887	6.928	6.6715	6.7966	6.7947
GaF_3	$R\bar{3}c$	5.208	5.208	5.208	5.2292	5.2292	5.2292
GaS	$P6_3/mmc$	3.585	3.585	15.500	3.6333	3.6333	15.6032
GaSe	$R3m$	3.755	3.755	15.940	3.8217	3.8217	16.1576
Ga_2O_3	$C2/m$	3.037	5.798	6.293	3.0748	5.8567	6.3773
Ga_2S_3	Cc	6.411	6.427	7.038	6.4605	6.4761	7.1033
Ga_2Se_3	Cc	6.660	6.660	6.710	6.7527	6.7435	6.8064
GeS	$Pmcn$	3.640	4.290	10.420	3.6926	4.4253	10.7480
GeSe	$Pmcn$	3.833	4.388	10.825	3.9133	4.4897	11.2631
HfO_2	$P2_1/c$	5.150	5.185	5.342	5.1269	5.1681	5.3128
$HgCl_2$	$Pcmn$	4.323	5.963	12.735	4.1607	6.0007	12.7550
HgO	$Pcmn$	3.521	5.520	6.612	3.5992	5.5529	6.7550
InS	$Pmnn$	3.940	4.443	10.642	3.9972	4.6967	10.6818
In_2O_3	$Ia\bar{3}$	8.781	8.781	8.781	8.8691	8.8691	8.8691
In_2S_3	$I4_1/amd$	7.617	7.617	17.039	7.7273	7.7273	17.3415
KSb	$P2_1/c$	6.879	7.072	12.037	6.9455	7.1587	12.1102
KSb_2	$C2/m$	4.233	7.053	7.339	4.3116	7.0467	7.4238
KO	$Ccme$	4.855	4.855	6.474	4.7841	4.7841	6.3953
KS	$P\bar{6}2m$	8.490	8.490	5.840	8.4002	8.4002	5.7075
K_3As	$P6_3/mmc$	5.794	5.794	10.242	5.7039	5.7039	10.1000
K_3Bi	$P6_3cm$	10.649	10.649	10.940	6.0661	6.0661	10.7669
K_3Sb	$P6_3/mmc$	6.037	6.037	10.715	5.9673	5.9673	10.5801
K_5Sb_4	$C2/m$	5.402	6.636	10.808	5.4074	6.6895	10.9740
$LaCl_3$	$P6_3/m$	7.483	7.483	4.375	7.4731	7.4731	4.3609
La_2O_3	$Ia\bar{3}$	9.890	9.890	9.890	9.8362	9.8362	9.8362
La_2S_3	$Pmnb$	4.220	7.660	15.950	4.1599	7.6230	15.8801
Li_2O	$Fm\bar{3}m$	3.382	3.382	3.382	3.2298	3.2298	3.2298
LiO	$P6_3/mmc$	3.183	3.183	7.726	3.1109	3.1109	7.5534
MgF_2	$P4_2/mnm$	3.096	4.691	4.691	3.0393	4.5935	4.5935
Mg_3Bi_2	$P\bar{3}m1$	4.666	4.666	7.401	4.6862	4.6861	7.4896
Mg_3Sb_2	$P\bar{3}m1$	4.559	4.559	7.227	4.5748	4.5748	7.2936

Table C.2 Experimental and calculated lattice constants of the compounds in the optimization set.

Comp.	Space group	Exp[127] a	Exp[127] b	Exp[127] c	PBEfe a	PBEfe b	PBEfe c
NaSb	$P2_1/c$	6.340	6.800	11.107	6.2893	6.7739	11.1220
$NaTe_3$	$P\bar{3}c1$	9.033	9.033	21.930	9.0520	9.0520	22.1166
NaO	$P\bar{6}2m$	6.207	6.207	4.471	6.1174	6.1174	4.3902
NaS	$P6_3/mmc$	4.494	4.494	10.228	4.4121	4.4121	10.0777
NaSe	$P6_3/mmc$	4.685	4.685	10.530	4.6207	4.6207	10.6461
Na_3As	$P6_3cm$	8.784	8.784	8.999	4.9988	4.9988	8.8975
Na_3Bi	$P6_3/mmc$	5.459	5.459	9.675	5.3770	5.3770	9.5601
Na_3Sb	$P6_3/mmc$	5.366	5.366	9.515	5.2746	5.2746	9.3789
NbO_2	$I4_1/a$	5.985	7.479	7.479	6.0835	7.5482	7.5482
NiF_2	$P4_2/mnm$	3.080	4.650	4.650	3.0894	4.6071	4.6071
NiSe	$R3m$	5.883	5.883	5.883	5.9047	5.9047	5.9047
$Ni3S_2$	$R32$	4.032	4.032	4.032	4.0839	4.0839	4.0839
PdS_2	$Pbca$	5.460	5.541	7.531	5.5520	5.6298	7.6431
$PdCl_2$	$R\bar{3}$	8.060	8.060	8.060	8.0471	8.0471	8.0471
PtO_2	$Pmnn$	3.138	4.488	4.533	3.1976	4.5963	4.5377
Pt_3O_4	$Pm\bar{3}n$	5.585	5.585	5.585	5.6639	5.6639	5.6639
RbSb	$P2_12_12_1$	7.197	7.315	12.815	7.2187	7.3200	12.7135
$RbSb_2$	$C2/m$	4.140	7.314	8.557	4.2372	7.2448	8.5540
Rb_3Sb	$P6_3/mmc$	6.283	6.283	11.180	6.2541	6.2541	11.0484
Sc_2O_3	$Ia\bar{3}$	8.418	8.418	8.418	8.5266	8.5266	8.5266
$ScCl_3$	$R\bar{3}$	6.979	6.979	6.979	6.8019	6.8019	6.8019
SiO_2	$I\bar{4}2d$	6.241	6.241	6.241	5.0792	5.0794	5.0795
SiS_2	$Imcb$	5.552	5.564	6.182	5.6023	5.5525	6.1656
$SiSe_2$	$Imcb$	5.851	5.998	6.397	5.9478	6.0022	6.3918
SnO_2	$P4_2/mnm$	3.195	4.708	4.720	3.2412	4.7991	4.8018
SnS	$Pmcn$	3.940	4.228	10.957	4.0277	4.4470	11.4424
SnSe	$Pmcn$	4.222	4.400	11.580	4.2118	4.5730	11.8303
Sr_2Bi	$I4/mmm$	5.010	5.010	9.523	5.0710	5.0710	9.6671
Sr_2Sb	$I4/mmm$	5.002	5.002	9.394	4.9924	4.9924	9.5173
TiAs	$P6_3/mmc$	3.642	3.642	12.064	3.6633	3.6633	12.0862
Ti_2O_3	$R\bar{3}c$	5.433	5.433	5.433	5.5020	5.5020	5.5020

Table C.2 Experimental and calculated lattice constants of the compounds in the optimization set.

	Space	Exp[127]			PBEfe		
Comp.	group	a	b	c	a	b	c
$TiCl_4$	$P2_1/c$	6.474	6.474	6.474	6.3070	9.4351	9.4390
TiO_2	$C2/m$	3.741	6.370	6.525	3.7403	6.3976	6.5959
VO_2	$P4_2/mnm$	4.526	5.381	5.381	4.5621	5.3793	5.3793
YCl_3	$C2/m$	6.440	6.900	6.900	6.2210	6.8598	6.8598
YF_3	$Pbnm$	4.393	6.353	6.850	4.3496	6.4125	6.8670
$ZnCl_2$	$I\bar{4}2d$	5.400	5.400	6.431	5.2439	5.2439	6.3824
ZnF_2	$P4_2/mnm$	3.131	4.715	4.715	3.1509	4.7020	4.7020
Zn_3As_2	$P4_2/nbc$	11.789	11.789	23.635	8.4467	8.4467	11.9343
Zn_3N_2	$Ia\bar{3}$	8.460	8.460	8.460	8.5145	8.5145	8.5145
ZnSb	$Pbca$	6.218	7.741	8.115	6.3379	7.7913	8.2158
ZrO_2	$P2_1/c$	5.164	5.164	5.164	5.2080	5.2614	5.3865
Al_2MgO_4	$Fd\bar{3}m$	5.733	5.733	5.733	5.7776	5.7776	5.7776
Al_2NiO_4	$Imcm$				5.7379	5.7379	5.7900
Al_2ZnO_4	$Fd\bar{3}m$	5.717	5.717	5.717	5.7827	5.7827	5.7827
Al_2ZnS_4	$Fd\bar{3}m$	7.048	7.048	7.048	7.1316	7.1316	7.1316
$BaTiO_3$	$R3m$	4.004	4.004	4.004	4.0764	4.0764	4.0764
Ba_2TiO_4	$P2_1/c$	6.120	7.700	10.500	6.2399	7.7746	10.6833
$BaCrO_4$	$Pmcn$	5.679	7.454	9.206	5.6865	7.4600	9.2084
$BaGeO_3$	$P2_12_12_1$	4.580	5.680	12.760	4.6807	5.8176	13.0985
$BaSi_2O_5$	$Pcmn$	4.629	7.688	13.523	4.7081	7.7919	13.7244
$BaSiO_3$	$C2/c$				7.5738	7.5738	11.4700
$BeAl_2O_4$	$Pcmn$	4.426	5.475	9.402	4.4615	5.5279	9.4999
Co_2SiO_4	$Pcmn$	4.779	5.987	10.286	4.8333	6.0640	10.3700
Cr_2FeO_4	$Fd\bar{3}m$	5.974	5.974	5.974	6.0424	6.0424	6.0424
Cr_2MgO_4	$Fd\bar{3}m$	5.966	5.966	5.966	5.9760	5.9760	5.9760
$CrMgO_4$	$Ccmm$	5.060	5.060	6.487	5.0607	5.0607	6.4796
Fe_2CdO_4	$Fd\bar{3}m$	6.244	6.244	6.244	6.2402	6.2402	6.2402
Fe_2CuO_4	$Fd\bar{3}m$				5.7727	5.7727	5.7729
Fe_2NiO_4	$Imma$	5.817	5.758	5.817	5.9512	5.9133	5.9512
Fe_2SiO_4	$Pcmn$	4.932	6.026	10.417	4.8844	6.1680	10.6243
Fe_2ZnO_4	$Fd\bar{3}m$	6.036	6.036	6.036	6.0292	6.0292	6.0292

Table C.2 Experimental and calculated lattice constants of the compounds in the optimization set.

Comp.	Space group	Exp[127] a	Exp[127] b	Exp[127] c	PBEfe a	PBEfe b	PBEfe c
$FeCuO_2$	$R\bar{3}m$	5.984	5.984	5.984	6.0450	6.0450	6.0450
$MgTi_2O_5$	$Cmcm$	3.745	5.212	9.990	3.7595	5.2603	10.1228
Mg_2SiO_4	$Pcmn$	4.765	5.997	10.230	4.8028	6.0479	10.3234
Mg_2TiO_4	$P4_122$				6.0517	6.0517	8.4876
$MgTiO_3$	$R\bar{3}$	5.477	5.477	5.477	5.5317	5.5317	5.5317
Sr_2SiO_4	$P2_1/c$	5.663	7.084	9.767	5.7542	7.1614	9.8536
Sr_2TiO_4	$I4/mmm$	3.884	3.884	6.873	3.9265	3.9265	6.9205
$SrAl_2O_4$	$P2_1$	5.173	8.507	8.890	5.2139	8.5474	8.9176
$SrSiO_3$	$C2/c$	7.127	7.127	10.885	7.2177	7.2177	10.9961
$SrTiO_3$	$I4/mcm$	5.511	5.511	5.512	5.5666	5.5666	5.5782
$TiZn_2O_4$	$P4_122$	6.005	6.005	8.416	6.0760	6.0760	8.5156
V_2MgO_6	$C2/m$	3.502	4.959	6.731	3.5171	5.3170	7.0782
$ZnTiO_3$	$R\bar{3}$	5.490	5.490	5.490	5.5498	5.5498	5.5498
$ZnSiO_3$	$C2/c$	5.289	6.694	6.694	5.3729	6.8868	6.8868

Table C.3 Experimental and calculated energies of formation of the compounds in the optimization set.

Comp.	Exp[159]	PBE	PBEsol	revPBE	xPBE	LDA	PBEfe	PBEfe+U
AgO	-0.06	-0.107	-0.082	-0.073	0.147	-0.138	-0.278	-0.188
AlAs	-0.61	-0.487	-0.433	-0.464	0.476	-0.465	-0.677	-0.677
AlN	-1.61	-1.409	-1.518	-1.261	1.324	-1.711	-1.800	-1.800
BaO	-2.86	-2.481	-2.509	-2.382	2.486	-2.636	-2.803	-2.803
BaO_2	-2.19	-1.853	-1.890	-1.756	1.857	-2.015	-2.133	-2.133
BaS	-2.38	-2.081	-2.094	-2.021	2.074	-2.141	-2.275	-2.275
BeO	-3.14	-2.762	-2.817	-2.633	2.740	-2.950	-3.157	-3.157
BeS	-1.21	-1.061	-1.096	-0.990	1.020	-1.146	-1.271	-1.271
CaF_2	-4.21	-3.900	-3.981	-3.823	3.966	-4.269	-4.293	-4.286
CaO	-3.29	-2.968	-3.023	-2.832	2.962	-3.214	-3.387	-3.387
CaS	-2.45	-2.155	-2.201	-2.061	2.126	-2.293	-2.412	-2.412
$CdCl_2$	-1.35	-1.097	-1.076	-1.070	1.110	-1.058	-1.129	-1.194
CdF_2	-2.42	-2.159	-2.200	-2.130	2.245	-2.407	-2.404	-2.494
CdO	-1.34	-1.031	-1.045	-0.963	1.034	-1.115	-1.209	-1.334
CdS	-0.78	-0.643	-0.621	-0.628	0.659	-0.619	-0.711	-0.769
CdSe	-0.75	-0.633	-0.569	-0.633	0.650	-0.540	-0.643	-0.692
CdTe	-0.48	-0.464	-0.384	-0.519	0.508	-0.362	-0.470	-0.504
CoO	-1.23	-0.672	-0.663	-0.666	0.752	-0.746	-0.947	-1.157
CoS	-0.43	-0.273	-0.414	-0.147	0.153	-0.516	-0.356	-0.355
CoSe	-0.32	-0.275	-0.302	-0.240	0.239	-0.347	-0.309	-0.258
CrN	-0.65	-0.343	-0.483	-0.231	0.239	-0.624	-0.524	-0.739
CrS	-0.81	-0.468	-0.508	-0.455	0.447	-0.555	-0.527	-0.786
CuO	-0.82	-0.600	-0.601	-0.540	0.621	-0.682	-0.821	-0.848
Cu_2Se	-0.21	0.013	0.021	0.036	-0.033	0.039	-0.007	-0.102
Cu_3N	0.19	0.282	0.255	0.326	-0.302	0.194	0.150	0.111
FeO	-1.41	-0.877	-0.746	-0.735	0.798	-0.738	-0.975	-1.360
FeS	-0.52	-0.515	-0.600	-0.439	0.453	-0.719	-0.632	-0.340
FeSe	-0.39	-0.281	-0.318	-0.224	0.229	-0.396	-0.343	-0.250
GaAs	-0.37	-0.352	-0.319	-0.335	0.343	-0.321	-0.371	-0.371
GaN	-0.81	-0.480	-0.643	-0.342	0.360	-0.781	-0.611	-0.611
GaSb	-0.22	-0.165	-0.123	-0.193	0.198	-0.125	-0.191	-0.191
HfN	-1.91	-1.752	-1.980	-1.560	1.559	-2.189	-1.973	-1.947

Table C.3 Experimental and calculated energies of formation of the compounds in the optimization set.

Comp.	Exp[159]	PBE	PBEsol	revPBE	xPBE	LDA	PBEfe	PBEfe+U
HgS	-0.3	-0.135	-0.179	0.019	0.029	-0.190	-0.159	-0.277
HgSe	-0.24	-0.215	-0.212	-0.080	0.118	-0.194	-0.183	-0.277
HgTe	-0.22	-0.171	-0.154	-0.088	0.098	-0.143	-0.137	-0.196
InAs	-0.31	-0.249	-0.203	-0.240	0.247	-0.209	-0.282	-0.282
InN	-0.1	0.082	-0.041	0.205	-0.173	-0.173	-0.065	-0.065
InSb	-0.16	-0.131	-0.081	-0.165	0.167	-0.091	-0.170	-0.170
InTe	-0.5	-0.298	-0.290	-0.295	0.286	-0.295	-0.329	-0.329
KCl	-2.26	-1.964	-1.943	-1.931	1.989	-1.992	-2.099	-2.099
KF	-2.94	-2.699	-2.738	-2.661	2.766	-2.964	-2.993	-2.987
K_2O	-1.25	-1.031	-1.054	-0.935	1.046	-1.221	-1.344	-1.344
K_2S	-1.31	-1.068	-1.072	-1.017	1.076	-1.153	-1.264	-1.264
K_2Se	-1.36	-1.110	-1.090	-1.073	1.114	-1.145	-1.255	-1.255
LaN	-1.57	-1.330	-1.391	-1.231	1.278	-1.506	-1.573	-1.588
LaS	-2.36	-2.170	-2.237	-2.070	2.097	-2.295	-2.323	-2.342
LiCl	-2.12	-1.816	-1.826	-1.747	1.817	-1.885	-1.991	-1.991
LiF	-3.19	-2.923	-2.598	-2.864	2.983	-3.221	-3.269	-3.263
Li_2O	-2.07	-1.837	-1.838	-1.743	1.858	-2.014	-2.187	-2.187
Li_2S	-1.52	-1.341	-1.342	-1.286	1.341	-1.423	-1.558	-1.558
Li_2Se	-1.45	-1.267	-1.251	-1.223	1.254	-1.307	-1.429	-1.429
Li_3Bi	-0.6	-0.535	-0.554	-0.511	0.500	-0.603	-0.626	-0.626
Li_3N	-0.43	-0.369	-0.407	-0.265	0.349	-0.565	-0.667	-0.667
Li_3Sb	-0.83	-0.637	-0.648	-0.617	0.608	-0.698	-0.734	-0.734
$MgCl_2$	-2.21	-1.864	-1.865	-1.809	1.860	-1.894	-2.013	-2.013
MgO	-3.11	-2.715	-2.753	-2.597	2.707	-2.933	-3.123	-3.123
MgS	-1.79	-1.433	-1.478	-1.355	1.385	-1.537	-1.642	-1.642
MgSe	-1.52	-1.257	-1.212	-1.227	1.253	-1.235	-1.428	-1.428
MgTe	-1.08	-0.879	-0.821	-0.908	0.893	-0.840	-1.021	-1.021
NaCl	-2.13	-1.813	-1.795	-1.775	1.834	-1.830	-1.951	-1.951
NaF	-2.97	-2.700	-2.727	-2.667	2.771	-2.938	-2.996	-2.990
Na_2O	-1.43	-1.219	-1.207	-1.149	1.250	-1.336	-1.517	-1.517
Na_2S	-1.26	-1.074	-1.063	-1.037	1.087	-1.119	-1.261	-1.261
Na_2Se	-1.18	-1.074	-1.043	-1.047	1.078	-1.076	-1.209	-1.209

Table C.3 Experimental and calculated energies of formation of the compounds in the optimization set.

Comp.	Exp[159]	PBE	PBEsol	revPBE	xPBE	LDA	PBEfe	PBEfe+U
NbN	-1.22	-1.077	-1.275	-0.906	0.911	-1.441	-1.263	-1.263
NiO	-1.24	-0.472	-0.608	-0.539	0.608	-0.661	-0.816	-1.306
NiSb	-0.34	-0.273	-0.323	-0.249	0.212	-0.338	-0.223	-0.242
PdO	-0.44	-0.445	-0.449	-0.381	0.471	-0.540	-0.664	-0.664
PtO	-0.37	-0.242	-0.231	-0.177	0.291	-0.338	-0.514	-0.514
PtS	-0.42	-0.411	-0.431	-0.355	0.401	-0.459	-0.518	-0.518
PtS_2	-0.38	-0.351	-0.406	-0.301	0.317	-0.455	-0.430	-0.430
RbCl	-2.26	-1.952	-1.933	-1.923	1.977	-1.967	-2.072	-2.072
RbF	-2.89	-2.639	-2.685	-2.601	2.702	-2.896	-2.921	-2.915
Rb_2O	-1.17	-0.919	-0.958	-0.811	0.919	-1.111	-1.216	-1.216
Rb_2S	-1.25	-1.009	-1.025	-0.953	1.009	-1.095	-1.193	-1.193
ScAs	-1.39	-1.375	-1.412	-1.307	1.280	-1.456	-1.457	-1.463
ScF_3	-4.22	-3.931	-3.989	-3.905	4.012	-4.216	-4.229	-4.265
SnO	-1.48	-1.290	-1.323	-1.218	1.300	-1.403	-1.458	-1.458
SnS_2	-0.53	-0.379	-0.429	-0.330	0.330	-0.437	-0.384	-0.384
$SnSe_2$	-0.43	-0.348	-0.354	-0.313	0.301	-0.335	-0.293	-0.293
SrO	-3.07	-2.739	-2.784	-2.620	2.737	-2.950	-3.116	-3.116
SrO_2	-2.19	-1.892	-1.932	-1.789	1.893	-2.078	-2.199	-2.199
SrS	-2.45	-2.151	-2.183	-2.071	2.130	-2.259	-2.382	-2.382
TiN	-1.58	-1.723	-1.941	-1.532	1.546	-2.145	-1.962	-1.918
TiS	-1.41	-1.427	-1.578	-1.296	1.287	-1.672	-1.525	-1.478
TiS_2	-1.41	-1.280	-1.333	-1.221	1.240	-1.389	-1.416	-1.486
VN	-1.13	-0.976	-1.181	-0.800	0.804	-1.362	-1.177	-1.073
YAs	-1.68	-1.553	-1.573	-1.496	1.470	-1.608	-1.645	-1.671
ZnO	-1.81	-1.445	-1.448	-1.382	1.475	-1.556	-1.692	-1.935
ZnS	-1.07	-0.811	-0.812	-0.769	0.817	-0.839	-0.929	-1.031
ZnSe	-0.85	-0.716	-0.673	-0.694	0.725	-0.667	-0.771	-0.849
ZnTe	-0.61	-0.470	-0.404	-0.506	0.508	-0.396	-0.514	-0.558
ZrN	-1.89	-1.683	-1.887	-1.503	1.512	-2.074	-1.914	-1.893
ZrS_2	-1.96	-1.519	-1.532	-1.480	1.503	-1.568	-1.663	-1.757

Table C.4 Experimental and calculated energies of formation of the compounds in the control set.

Comp.	Exp[159]	PBE	LDA	PBEfe	PBEfeU
Ag_2O	-0.11	-0.096	-0.028	-0.195	-0.252
Ag_2S	-0.11	-0.062	-0.010	-0.068	-0.159
Ag_2Se	-0.15	-0.079	-0.033	-0.046	-0.185
$AlCl_3$	-1.82	-1.510	-1.558	-1.643	-1.634
AlF_3	-3.9	-3.552	-3.758	-3.906	-3.900
Al_2O_3	-3.47	-3.017	-3.279	-3.407	-3.442
Al_2S_3	-1.5	-1.061	-1.107	-1.270	-1.276
Al_2Se_3	-1.18	-0.846	-0.817	-0.982	-0.994
Al_2Te_3	-0.68	-0.451	-0.380	-0.537	-0.548
$Be3N_2$	-1.22	-1.078	-1.361	-1.337	-1.383
$CaCl_2$	-2.75	-2.403	-2.420	-2.568	-2.558
CdSb	-0.07	-0.051	-0.006	-0.045	-0.086
Cd_3As_2	-0.08	-0.103	-0.023	-0.090	-0.149
Cd_3N_2	-0.33	0.549	0.396	0.463	0.350
CoF_2	-2.39	-1.702	-1.666	-1.928	-2.172
CoF_3	-2.1	-1.712	-1.865	-1.924	-2.035
$CoSb_3$	-0.17	-0.192	-0.265	-0.114	-0.144
Co_3O_4	-1.32	-0.987	-1.137	-1.216	-1.393
Co_3S_4	-0.53	-0.430	-0.655	-0.505	-0.541
CrO_2	-2.07	-1.900	-1.964	-2.161	-2.306
Cr_2O_3	-2.36	-1.926	-1.900	-2.163	-2.435
CuF_2	-1.88	-1.517	-1.610	-1.733	-1.855
CuS	-0.28	-0.199	-0.308	-0.279	-0.324
CuSe	-0.21	-0.137	-0.156	-0.187	-0.211
Cu_2O	-0.58	-0.414	-0.479	-0.575	-0.656
Cu_2Sb	-0.04	0.011	-0.047	0.108	0.050
Cu_2Te	0.07	0.086	0.096	0.108	0.043
Cu_3Sb	-0.02	0.067	0.015	0.163	0.105
FeF_2	-2.46	-1.916	-1.888	-2.138	-2.506
Fe_2O_3	-1.71	-1.172	-1.460	-1.451	-1.589
Fe_3O_4	-1.66	-1.122	-1.319	-1.398	-1.632
$GaCl_3$	-1.36	-1.111	-1.047	-1.101	-1.072

Table C.4 Experimental and calculated energies of formation of the compounds in the control set.

Comp.	Exp[159]	PBE	LDA	PBEfe	PBEfeU
GaF_3	-3.01	-2.609	-2.839	-2.802	-2.794
GaS	-1.09	-0.652	-0.689	-0.692	-0.708
GaSe	-0.83	-0.595	-0.571	-0.575	-0.596
Ga_2O_3	-2.26	-1.858	-2.081	-2.025	-2.053
Ga_2S_3	-1.07	-0.664	-0.707	-0.715	-0.724
Ga_2Se_3	-0.85	-0.589	-0.572	-0.575	-0.592
GeS	-0.39	-0.333	-0.339	-0.333	-0.338
GeSe	-0.48	-0.249	-0.245	-0.249	-0.246
HfO2	-3.95	-3.566	-3.827	-3.892	-3.949
$HgCl_2$	-0.77	-0.605	-0.567	-0.581	-0.691
HgO	-0.47	-0.294	-0.370	-0.385	-0.527
InS	-0.7	-0.509	-0.580	-0.583	-0.600
In_2O_3	-1.92	-1.600	-1.787	-1.778	-1.799
In_2S_3	-0.74	-0.557	-0.609	-0.621	-0.621
KSb	-0.43	-0.443	-0.465	-0.536	-0.542
KSb_2	-0.37	-0.304	-0.313	-0.359	-0.371
KO	-1.28	-1.066	-1.221	-1.334	-1.332
KS	-1.12	-0.940	-1.013	-1.112	-1.101
K_3As	-0.48	-0.335	-0.392	-0.490	-0.490
K_3Bi	-0.6	-0.382	-0.438	-0.511	-0.514
K_3Sb	-0.47	-0.418	-0.469	-0.556	-0.555
K_5Sb_4	-0.44	-0.444	-0.469	-0.549	-0.553
$LaCl_3$	-2.78	-2.414	-2.465	-2.558	-2.578
La_2O_3	-3.72	-3.468	-3.571	-3.780	-3.805
La_2S_3	-2.51	-2.201	-2.279	-2.377	-2.404
Li_2O	-2.07	-1.838	-2.017	-2.161	-2.188
Li_2O_2	-1.64	-1.416	-1.565	-1.697	-1.712
MgF_2	-3.88	-3.550	-3.754	-3.896	-3.889
Mg_3Bi_2	-0.32	-0.208	-0.265	-0.277	-0.310
Mg_3Sb_2	-0.49	-0.370	-0.414	-0.448	-0.474
NaSb	-0.33	-0.333	-0.341	-0.383	-0.404
$NaTe_3$	-0.35	-0.388	-0.383	-0.417	-0.429

Table C.4 Experimental and calculated energies of formation of the compounds in the control set.

Comp.	Exp[159]	PBE	LDA	PBEfe	PBEfeU
NaO	-1.32	-1.077	-1.231	-1.327	-1.338
NaS	-1.03	-0.833	-0.898	-0.982	-0.985
NaSe	-0.97	-0.834	-0.835	-0.918	-0.926
Na_3As	-0.53	-0.423	-0.454	-0.533	-0.551
Na_3Bi	-0.46	-0.384	-0.421	-0.468	-0.490
Na_3Sb	-0.53	-0.441	-0.471	-0.534	-0.552
NbO_2	-2.75	-2.447	-2.598	-2.785	-2.785
NiF_2	-2.25	-1.656	-1.677	-1.879	-2.311
NiSe	-0.31	-0.287	-0.396	-0.255	-0.338
Ni_3S_2	-0.42	-0.374	-0.562	-0.388	-0.454
$PdCl_2$	-0.69	-0.482	-0.474	-0.521	-0.521
PdS_2	-0.28	-0.268	-0.357	-0.317	-0.342
PtO_2	-0.57	-0.465	-0.627	-0.681	-0.717
$Pt3O_4$	-0.4	-0.440	-0.563	-0.647	-0.689
RbSb	-0.52	-0.433	-0.452	-0.525	-0.525
$RbSb_2$	-0.35	-0.304	-0.289	-0.356	-0.363
Rb_3Sb	-0.45	-0.365	-0.419	-0.493	-0.493
$ScCl_3$	-2.4	-2.125	-2.127	-2.222	-2.265
Sc_2O_3	-3.94	-3.559	-3.763	-3.925	-3.947
SiO_2	-3.13	-2.980	-2.932	-2.980	-2.996
SiS_2	-0.88	-0.657	-0.657	-0.657	-0.654
$SiSe_2$	-0.61	-0.376	-0.370	-0.376	-0.366
SnO_2	-1.97	-1.652	-1.888	-1.805	-1.823
SnS	-0.57	-0.443	-0.543	-0.500	-0.515
SnSe	-0.47	-0.436	-0.474	-0.426	-0.446
Sr_2Bi	-1.08	-0.761	-0.834	-0.829	-0.864
Sr_2Sb	-1.11	-0.864	-0.927	-0.941	-0.970
TiAs	-0.78	-0.926	-1.008	-0.887	-0.920
$TiCl_4$	-1.7	-1.597	-1.456	-1.616	-1.626
TiO_2	-3.26	-3.048	-3.092	-3.358	-3.402
Ti_2O_3	-3.15	-2.896	-3.078	-3.202	-3.259
VO_2	-2.47	-2.346	-2.467	-2.622	-2.763

Table C.4 Experimental and calculated energies of formation of the compounds in the control set.

Comp.	Exp[159]	PBE	LDA	PBEfe	PBEfeU
YCl_3	-2.59	-2.290	-2.245	-2.397	-2.420
YF_3	-4.45	-4.112	-4.332	-4.450	-4.511
$ZnCl_2$	-1.43	-1.149	-1.101	-1.201	-1.296
ZnF_2	-2.64	-2.318	-2.432	-2.547	-2.738
ZnSb	-0.08	-0.032	0.004	-0.068	-0.085
$Zn3As_2$	-0.28	-0.144	-0.084	-0.185	-0.241
Zn_3N_2	-0.05	0.114	-0.106	-0.040	-0.250
ZrO2	-3.8	-3.360	-3.554	-3.688	-3.826
Al_2MgO_4	-3.4	-3.403	-3.222	-2.985	-3.403
Al_2NiO_4	-2.82	-2.681	-2.492	-2.311	-2.841
Al_2ZnO_4	-3.02	-3.004	-2.850	-2.630	-3.075
Al_2ZnS_4	-1.36	-1.189	-1.112	-1.017	-1.217
$BaGeO_3$	-2.6	-2.466	-2.424	-2.286	-2.645
$BaTiO_3$	-3.35	-3.421	-2.367	-3.083	-2.466
Ba_2TiO_4	-3.18	-3.315	-3.132	-3.010	-3.235
$BaCrO_4$	-2.42	-2.650	-3.074	-2.419	-3.163
$BaSi_2O_5$	-3.21	-3.163	-3.256	-2.937	-3.456
$BaSiO_3$	-3.27	-3.235	-3.089	-2.997	-3.296
$BeAl_2O_4$	-3.34	-3.377	-3.205	-2.968	-3.377
Co_2SiO_4	-2.1	-1.722	-1.556	-1.501	-1.902
$CrMgO_4$	-2.17	-2.409	-2.169	-2.151	-2.404
Cr_2FeO_4	-2.19	-1.739	-1.591	-1.645	-2.168
Cr_2MgO_4	-2.65	-2.514	-2.229	-2.204	-2.689
$FeCuO_2$	-1.3	-1.196	-1.121	-0.902	-1.317
Fe_2CdO_4	-1.59	-1.336	-1.279	-1.336	-1.580
Fe_2CuO_4	-1.43	-1.312	-1.294	-1.076	-1.393
Fe_2NiO_4	-1.6	-1.237	-1.222	-0.979	-1.539
Fe_2SiO_4	-2.23	-1.871	-1.696	-1.655	-2.152
Fe_2ZnO_4	-1.73	-1.563	-1.475	-1.240	-1.755
$MgTiO_3$	-3.24	-3.319	-3.109	-2.962	-3.348
$MgTi_2O_5$	-3.25	-3.339	-3.102	-3.002	-3.378
Mg_2SiO_4	-3.17	-3.154	-3.040	-2.838	-3.154

Table C.4 Experimental and calculated energies of formation of the compounds in the control set.

Comp.	Exp[159]	PBE	LDA	PBEfe	PBEfeU
Mg_2TiO_4	-3.21	-3.265	-3.055	-2.898	-3.295
$SrTiO_4$	-3.42	-3.502	-3.222	-3.145	-3.470
Sr_2SiO_4	-3.35	-3.337	-3.193	-3.039	-3.296
Sr_2TiO_3	-3.36	-3.415	-3.324	-3.057	-3.538
$SrAl_2O_4$	-3.47	-3.470	-3.236	-3.074	-3.337
$SrSiO_3$	-3.33	-3.296	-3.245	-3.038	-3.438
$TiZn_2O_4$	-2.38	-2.381	-2.217	-2.108	-2.554
$V2MgO_6$	-2.53	-2.717	-2.498	-2.448	-2.460
$ZnTiO_3$	-2.63	-2.674	-2.382	-2.382	-2.564
$ZnSiO_3$	-2.52	-2.464	-2.498	-2.236	-2.798

Table C.5 Experimental lattices constants of the elementary solids from references [172] and [127](∗).

Comp.	Spg.	a	b	c
Fe	229	2.8665	2.8665	2.8665
La	194	3.770	3.770	12.159
Ni	225	3.523	3.523	3.523
Ga	64	4.5167	7.6448	4.5107
Ag	225	4.08626	4.08626	4.08626
Ba	229	5.013	5.013	5.013
Ca	225	5.5884	5.5884	5.5884
Co	194	2.5071	2.5071	4.0695
Cu	225	3.61491	3.61491	3.61491
Hg	139	3.995	3.995	2.825
K	229	5.148	5.148	5.148
Li(∗)	166	3.012	3.012	3.012
Nb	229	3.3063	3.3063	3.3063
Pd	225	3.8874	3.8874	3.8874
Rb	229	5.585	5.585	5.585
Sr	225	6.0849	6.0849	6.0849
Ti	194	2.9503	2.9503	4.6810
Y	194	3.6515	3.6515	5.7474
Zr	194	3.23178	3.23178	5.14831
Al	225	4.04950	4.04950	4.04950
Be	194	2.2858	2.2858	3.5843
Cd	194	2.972	2.972	5.605
Cr	229	2.8844	2.8844	2.8844
Hf	194	3.198	3.198	5.061
In	139	3.2530	3.2530	4.9455
Mg	194	3.20944	3.20944	5.21076
Na	166	3.767	3.767	6.154
Pt	225	3.924	3.924	3.924
Sc	194	3.309	3.309	5.2733
Sn	227	3.287	3.287	3.287
V	229	3.0309	3.0309	3.0309
Zn	194	2.66469	2.66469	4.94616

Table C.5 Experimental lattices constants of the elementary solids from references [172] and [127](∗).

Comp.	Spg.	a	b	c
H	194	3.776	3.776	6.162
Bi	166	4.5460	4.5460	11.862
C	194	2.464	2.464	6.711
Se	152	4.366	4.366	4.955
N	205	6.164	6.164	6.164
B	166	4.91	4.91	12.57
F	15	5.50	3.38	7.28
O(∗)	12	3.773	4.983	4.983
Si	227	3.34	3.34	3.34
I	64	7.2697	4.7903	9.7942
Br	64	6.68	4.49	8.74
Te	152	4.456	4.456	5.921
Cl	64	6.24	4.48	8.26
Ge	227	5.65752	5.65752	5.65752

Appendix D

Experimental Absorption Edges Reproduced from Johnson *et al.* [131]

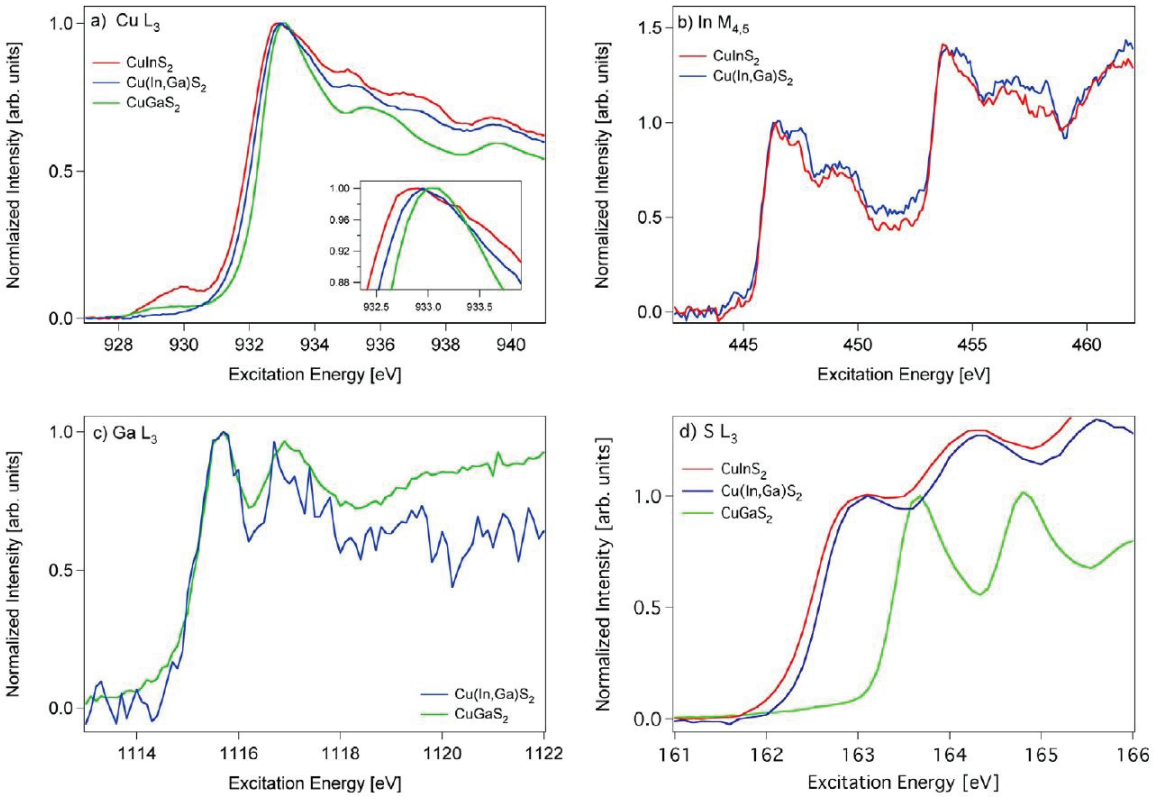


Fig. D.1 Absorption edges reproduced from Reference [131] of (a) Cu L_3 , (b) In $M_{4,5}$, (c) Ga L_3 and (d) S L_3 from $CuInS_2$ (red), $CuIn_{0.67}Ga_{0.33}S_2$ (blue) and $CuGaS_2$ (green) measured with total electron yield. The energy scale for each set of curves is not absolute, but the edge positions are correct relative to one another, making shifts in the spectra correspond to real energy differences between the core level and the end state of the electron in the conduction bands of the three materials. Inset: the crests of the Cu L_3 spectra.